PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, 15/54, 15/55, 9/10, 9/14, 5/10, C12Q 1/68, G01N 33/50, A01H 5/00 — A2

(11) International Publication Number: WO 99/38986

(43) International Publication Date: 5 August 1999 (05.08.99)

(21) International Application Number: PCT/EP99/00556

(22) International Filing Date: 28 January 1999 (28.01.99)

(30) Priority Data:
60/109,810 30 January 1998 (30.01.98) US

(71) Applicant *(for all designated States except AT US)*: NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH–4058 Basel (CH).

(71) Applicant *(for AT only)*: NOVARTIS–ERFINDUNGEN VERWALTUNGSGESELLSCHAFT M.B.H. [AT/AT]; Brunner Strasse 59, A–1235 Vienna (AT).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: GUYER, Charles, David [CA/US]; 4615 J Hope Valley Road, Durham, NC 27707 (US). JOHNSON, Marie, Ann [US/US]; 408 Heather Drive, Raleigh, NC 27606 (US). VOLRATH, Sandra, Lynn [US/US]; 4225 Pine Oak Drive, Durham, NC 27707 (US). BRUNN, Sandra, Alice [CA/US]; 523 Sioux Lane, San Jose, CA 95123 (US). WARD, Eric, Russell [US/US]; 3761 Bentley Drive, Durham, NC 27707 (US).

(74) Agent: BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent & Trademark Dept., CH–4002 Basel (CH).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: RIBOFLAVIN BIOSYNTHESIS GENES FROM PLANTS AND USES THEREOF

(57) Abstract

The present invention provides plant riboflavin biosynthesis genes, including a gene that encodes the β subunit of the plant riboflavin synthase enzyme complex (lumazine synthase) and a gene that encodes the bifunctional enzyme GTP cyclohydrolase II/DHBP synthase. Also disclosed are the recombinant production of these plant riboflavin biosynthesis enzymes in heterologous hosts, screening chemicals for herbicidal activity using these recombinantly produced enzymes, and the use of thereby identified herbicidal chemicals to suppress the growth of undesired vegetation. Furthermore, the present invention provides methods for the development of herbicide tolerance in plants, plant tissues, plant seeds and plant cells using the riboflavin biosynthesis genes of the invention.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

RIBOFLAVIN BIOSYNTHESIS GENES FROM PLANTS AND USES THEREOF

The invention relates generally to enzymatic activity involved in riboflavin biosynthesis in plants. In particular, the invention relates to plant genes that encode the bifunctional GTP cyclohydrolase II / DHBP synthase enzyme and the β subunit of the riboflavin synthase enzyme complex (lumazine synthase). The invention has various utilities, including the recombinant production of these riboflavin biosynthesis enzymes in heterologous hosts, the screening of chemicals for herbicidal activity, and the use of thereby identified herbicidal chemicals to control the growth of undesired vegetation. The invention may also be applied to the development of herbicide tolerance in plants, plant tissues, plant seeds, and plant cells.

I. Riboflavin Biosynthesis

Riboflavin (vitamin B_2 – 6,7-dimethyl-9-(1-D-ribityl)-isoalloxazine) is synthesized by all plants and many microorganisms. Riboflavin is essential to basic metabolism because it is a precursor to coenzymes such as FAD and FMN, which are required in the enzymatic oxidation of carbohydrates. Biosynthesis of riboflavin starts from guanosine-5'-triphosphate (GTP) and proceeds through several enzymatic steps, as outlined in Figure 1 of Mironov *et al.*, *Mol. Gen. Genet.* 242:201-208 (1994), incorporated herein by reference.

GTP cyclohydrolase II is the first enzyme of riboflavin biosynthesis, catalyzing the synthesis of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate from GTP. DHBP synthase catalyzes the conversion of ribulose-5-phosphate to 3,4-dihydroxy-2-butanone phosphate (DHBP). In *Bacillus*, these two enzymatic activities are carried out by a single, bifunctional enzyme; in *E. coli*, however, these two enzymatic activities are carried out by two separate enzymes.

The riboflavin synthase protein is an approximately 1,000,000-Da enzyme complex consisting of approximately 60 β subunits and three α subunits. The β subunits form a capsid that catalyzes the conversion of 2,4-dioxy-5-amino-6-ribitylamino-pyrimidine (DARP) and 3,4-dihydroxy-2-butanone phosphate (DHBP) to 6,7-dimethyl-8-ribityllumazine (lumazine); hence, the β subunit is also known as "lumazine synthase". The α subunits, contained inside the β subunit capsid, then catalyze the conversion of two units of lumazine to one DARP molecule, which is recycled back into the first riboflavin synthase reaction, and one riboflavin molecule.

II. Herbicide Discovery

The use of herbicides to control undesirable vegetation such as weeds in crop fields has become almost a universal practice. The herbicide market exceeds 15 billion dollars annually. Despite this extensive use, weed control remains a significant and costly problem for farmers.

Effective use of herbicides requires sound management. For instance, the time and method of application and stage of weed plant development are critical to getting good weed control with herbicides. Since various weed species are resistant to herbicides, the production of effective new herbicides becomes increasingly important. Novel herbicides can now be discovered using high-throughput screens that implement recombinant DNA technology. Metabolic enzymes essential to plant growth and development can be recombinantly produced though standard molecular biological techniques and utilized as herbicide targets in screens for novel inhibitors of the enzymes' activity. The novel inhibitors discovered through such screens may then be used as herbicides to control undesirable vegetation.

III. Herbicide Tolerant Plants

Herbicides that exhibit greater potency, broader weed spectrum, and more rapid degradation in soil can also, unfortunately, have greater crop phytotoxicity. One solution applied to this problem has been to develop crops that are resistant or tolerant to herbicides. Crop hybrids or varieties tolerant to the herbicides allow for the use of the herbicides to kill weeds without attendant risk of damage to the crop. Development of tolerance can allow application of a herbicide to a crop where its use was previously precluded or limited (*e.g.* to pre-emergence use) due to sensitivity of the crop to the herbicide. For example, U.S. Patent No. 4,761,373 to Anderson *et al.* is directed to plants resistant to various imidazolinone or sulfonamide herbicides. The resistance is conferred by an altered acetohydroxyacid synthase (AHAS) enzyme. U.S. Patent No. 4,975,374 to Goodman *et al.* relates to plant cells and plants containing a gene encoding a mutant glutamine synthetase (GS) resistant to inhibition by herbicides that were known to inhibit GS, *e.g.* phosphinothricin and methionine sulfoximine. U.S. Patent No. 5,013,659 to Bedbrook *et al.* is directed to plants expressing a mutant acetolactate synthase that renders the plants resistant to inhibition by sulfonylurea herbicides. U.S. Patent No. 5,162,602 to Somers *et al.* discloses plants tolerant to inhibition by cyclohexanedione and

aryloxyphenoxypropanoic acid herbicides. The tolerance is conferred by an altered acetyl coenzyme A carboxylase (ACCase).

DEFINITIONS

For clarity, certain terms used in the specification are defined and presented as follows:

Activatable DNA Sequence: a DNA sequence that regulates the expression of genes in a genome, desirably the genome of a plant. The activatable DNA sequence is complementary to a target gene endogenous in the genome. When the activatable DNA sequence is introduced and expressed in a cell, it inhibits expression of the target gene. An activatable DNA sequence useful in conjunction with the present invention includes those encoding or acting as dominant inhibitors, such as a translatable or untranslatable sense sequence capable of disrupting gene function in stably transformed plants to positively identify one or more genes essential for normal growth and development of a plant. A preferred activatable DNA sequence is an antisense DNA sequence. The target gene preferably encodes a protein, such as a biosynthetic enzyme, receptor, signal transduction protein, structural gene product, or transport protein that is essential to the growth or survival of the plant. In an especially preferred embodiment, the target gene encodes lumazine synthase or the bifunctional enzyme GTP cyclohydrolase II / DHBP synthase. The interaction of the antisense sequence and the target gene results in substantial inhibition of the expression of the target gene so as to kill the plant, or at least inhibit normal plant growth or development.

Activatable DNA Construct: a recombinant DNA construct comprising a synthetic promoter operatively linked to the activatable DNA sequence, which when introduced into a cell, desirably a plant cell, is not expressed, i.e. is silent, unless a complete hybrid transcription factor capable of binding to and activating the synthetic promoter is present. The activatable DNA construct is introduced into cells, tissues, or plants to form stable transgenic lines capable of expressing the activatable DNA sequence.

Chimeric: "chimeric" is used to indicate that a DNA sequence, such as a vector or a gene, is comprised of more than one DNA sequences of distinct origin which are fused together by recombinant DNA techniques resulting in a DNA sequence, which does not occur naturally, and which particularly does not occur in the plant to be transformed.

DNA shuffling: DNA shuffling is a method to introduce mutations or rearrangements, preferably randomly, in a DNA molecule or to generate exchanges of DNA sequences between two or more DNA molecules, preferably randomly. The DNA molecule resulting from DNA shuffling is a shuffled DNA molecule that is a non-naturally occurring DNA molecule derived from at least one template DNA molecule. The shuffled DNA encodes an enzyme modified with respect to the enzyme encoded by the template DNA, and preferably has an altered biological activity with respect to the enzyme encoded by the template DNA.

Enzyme activity: means herein the ability of an enzyme to catalyze the conversion of a substrate into a product. A substrate for the enzyme comprises the natural substrate of the enzyme but also comprises analogues of the natural substrate which can also be converted by the enzyme into a product or into an analogue of a product. The activity of the enzyme is measured for example by determining the amount of product in the reaction after a certain period of time, or by determining the amount of substrate remaining in the reaction mixture after a certain period of time. The activity of the enzyme is also measured by determining the amount of an unused co-factor of the reaction remaining in the reaction mixture after a certain period of time or by determining the amount of used co-factor in the reaction mixture after a certain period of time. The activity of the enzyme is also measured by determining the amount of a donor of free energy or energy-rich molecule (e.g. ATP, phosphoenolpyruvate, acetyl phosphate or phosphocreatine) remaining in the reaction mixture after a certain period of time or by determining the amount of a used donor of free energy or energy-rich molecule (e.g. ADP, pyruvate, acetate or creatine) in the reaction mixture after a certain period of time.

Expression refers to the transcription and/or translation of an endogenous gene or a transgene in plants. In the case of antisense constructs, for example, expression may refer to the transcription of the antisense DNA only.

Gene refers to a coding sequence and associated regulatory sequences wherein the coding sequence is transcribed into RNA such as mRNA, rRNA, tRNA, snRNA, sense RNA or antisense RNA. Examples of regulatory sequences are promoter sequences, 5' and 3' untranslated sequences and

Herbicide: a chemical substance used to kill or suppress the growth of plants, plant cells, plant seeds, or plant tissues.

Heterolog us DNA Sequence: a DNA sequence not naturally associated with a host cell into which it is introduced, including non-naturally occurring multiple copies of a naturally occurring DNA sequence.

Homologous DNA Sequence: a DNA sequence naturally associated with a host cell into which it is introduced.

Inhibitor: a chemical substance that inactivates the enzymatic activity of a protein such as a biosynthetic enzyme, receptor, signal transduction protein, structural gene product, or transport protein that is essential to the growth or survival of the plant. In the context of the instant invention, an inhibitor is a chemical substance that inactivates the enzymatic activity of lumazine synthase or the bifunctional enzyme GTP cyclohydrolase II / DHBP synthase from a plant. The term "herbicide" is used herein to define an inhibitor when applied to plants, plant cells, plant seeds, or plant tissues.

Isolated: in the context of the present invention, an isolated DNA molecule or an isolated enzyme is a DNA molecule or enzyme, by the hand of man, exists apart from its native environment and is therefore not a product of nature. An isolated DNA molecule or enzyme may exist in a purified form or may exist in a non-native environment such as, for example, a transgenic host cell.

Minimal Promoter: promoter elements, particularly a TATA element, that are inactive or that have greatly reduced promoter activity in the absence of upstream activation. In the presence of a suitable transcription factor, the minimal promoter functions to permit transcription.

Modified Enzyme Activity: enzyme activity different from that which naturally occurs in a plant (i.e. enzyme activity that occurs naturally in the absence of direct or indirect manipulation of such activity by man), which is tolerant to inhibitors that inhibit the naturally occurring enzyme activity.

Plant refers to any plant, particularly to seed plants

Plant cell: structural and physiological unit of the plant, comprising a protoplast and a cell wall. The plant cell may be in form of an isolated single cell or a cultured cell, or as a part of higher organized unit such as, for example, a plant tissue, or a plant organ.

Recombinant DNA: molecule a combination of DNA sequences that are joined together using recombinant DNA technology

Recombinant DNA technology: procedures used to join together DNA sequences as described, for example, in Sambrook et al., 1989, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Significant increase: an increase in enzymatic activity that is larger than the margin of error inherent in the measurement technique, preferably an increase by about 2-fold or greater of the activity of the wild-type enzyme in the presence of the inhibitor, more preferably an increase by about 5-fold or greater, and most preferably an increase by about 10-fold or greater.

Significantly less: means that the amount of a product of an enzymatic reaction is larger than the margin of error inherent in the measurement technique, preferably a decrease by about 2-fold or greater of the activity of the wild-type enzyme in the absence of the inhibitor, more preferably an decrease by about 5-fold or greater, and most preferably an decrease by about 10-fold or greater.

Substantially Similar: in the context of the present invention, a DNA molecule that has at least 60 percent sequence identity with the portion of SEQ ID NO:1 that codes for lumazine synthase, i.e. that portion of SEQ ID NO:1 that encodes the amino acid sequence of SEQ ID NO:2; or a DNA molecule that has at least 60 percent sequence identity with the portion of SEQ ID NO:13 that codes for the bifunctional GTP cyclohydrolase II / DHBP synthase enzyme from a plant, i.e. that portion of SEQ ID NO:13 that encodes the amino acid sequence of SEQ ID NO:14. A substantially similar lumazine synthase nucleotide sequence hybridizes specifically to SEQ ID NO:1 or fragments thereof under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C. A substantially similar plant GTP cyclohydrolase II / DHBP synthase nucleotide sequence hybridizes specifically to SEQ ID NO:13 or fragments thereof under the above conditions. With respect to proteins, "substantially similar" as used herein means a protein sequence that is at least 90% identical to either the amino acid sequence set forth in SEQ ID NO:2 or the amino acid sequence set forth in SEQ ID NO:14.

Substrate: a substrate is the molecule that the enzyme naturally recognizes and converts to a product in the biochemical pathway in which the enzyme naturally carries out its function, or is a modified version of the molecule, which is also recognized by the enzyme and is converted by the enzyme to a product in an enzymatic reaction similar to the naturally-occurring reaction.

Synthetic refers to a nucleotide sequence comprising structural characters that are not present in the natural sequence. For example, an artificial sequence that resembles more closely the G+C content and the normal codon distribution of dicot and/or monocot genes is said to be synthetic.

Tolerance; the ability to continue normal growth or function when exposed to an inhibitor or herbicide.

Transformation: a process for introducing heterologous DNA into a cell, tissue, or plant. Transformed cells, tissues, or plants are understood to encompass not only the end product of a transformation process, but also transgenic progeny thereof.

Transgenic: stably transformed with a recombinant DNA molecule that preferably comprises a suitable promoter operatively linked to a DNA sequence of interest.

In view of the above, one object of the invention is to provide methods for identifying new or improved herbicides. Another object of the invention is to provide methods for using such new or improved herbicides to suppress the growth of plants such as weeds. Still another object of the invention is to provide improved crop plants that are tolerant to such new or improved herbicides.

In furtherance of these and other objects, the present invention provides a DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme involved in riboflavin biosynthesis, wherein the enzyme has either lumazine synthase activity or bifunctional GTP cyclohydrolase II / DHBP synthase activity.

According to one embodiment, the present invention provides a DNA molecule comprising a nucleotide sequence isolated from a plant that encodes the β subunit of riboflavin synthase (lumazine synthase). For example, the DNA molecule of the invention may comprises a nucleotide sequence that encodes an enzyme having lumazine synthase activity, wherein the enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:2. In another example, the DNA molecule of the invention comprises a nucleotide sequence that encodes an enzyme having lumazine synthase activity, wherein the enzyme comprises the amino acid sequence set forth in SEQ ID NO:2. In an other example, the DNA molecule of the invention comprises a nucleotide sequence isolated from a plant that encodes an enzyme having lumazine synthase activity, wherein said DNA molecule hybridizes to a DNA molecule that encodes the amino acid

sequence set forth in SEQ ID NO:2 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C. The invention further provides a DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme involved in riboflavin biosynthesis, wherein the enzyme has lumazine synthase activity, wherein said DNA molecule hybridizes to the coding sequence set forth in SEQ ID NO:1 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C. In yet another example, the DNA molecule of the invention comprises a nucleotide sequence that is substantially similar to the coding sequence set forth in SEQ ID NO:1 and that encodes an enzyme having lumazine synthase activity. In a further example, the DNA molecule of the invention comprises a nucleotide sequence isolated from a plant that encodes an enzyme having lumazine synthase activity, wherein said DNA molecule comprises a 20 base pair nucleotide portion identical in sequence to a consecutive 20 base pair portion of the coding sequence set forth in SEQ ID NO:1. In still another example, the DNA molecule of the invention comprises the coding sequence set forth in SEQ ID NO:1 and encodes an enzyme having lumazine synthase activity. Although the nucleotide sequence provided in SEQ ID NO:1 that encodes lumazine synthase was isolated from *Arabidopsis thaliana*, using the information provided by the present invention, the nucleotide sequence that encodes an enzyme having lumazine synthase activity can be obtained from any plant using standard methods known in the art.

According to another embodiment, the present invention provides a DNA molecule comprising a nucleotide sequence isolated from a plant that encodes the bifunctional GTP cyclohydrolase II / DHBP synthase. For example, the DNA molecule of the invention may comprise a nucleotide sequence that encodes an enzyme having bifunctional GTP cyclohydrolase II / DHBP synthase activity, wherein the enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:14. In another example, the DNA molecule of the invention comprises a nucleotide sequence that encodes an enzyme having bifunctional GTP cyclohydrolase II / DHBP synthase activity, wherein the enzyme comprises the amino acid sequence set forth in SEQ ID NO:14. In another example of the invention the DNA molecule comprises a nucleotide sequence isolated from a plant that encodes an enzyme having bifunctional GTP cyclohydrolase II / DHBP synthase activity, wherein said DNA molecule hybridizes to a DNA molecule that

encodes the amino acid sequence set forth in SEQ ID NO: 14 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C. The invention further provides a DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme involved in riboflavin biosynthesis, wherein the enzyme has bifunctional GTP cyclohydrolase II / DHBP synthase activity, wherein said DNA molecule hybridizes to the coding sequence set forth in SEQ ID NO:13 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C.

In yet another example, the DNA molecule of the invention comprises a nucleotide sequence that is substantially similar to the coding sequence set forth in SEQ ID NO:13 and that encodes an enzyme having bifunctional GTP cyclohydrolase II / DHBP synthase activity. In a further example, the DNA molecule of the invention comprises a nucleotide sequence isolated from a plant that encodes an enzyme having bifunctional GTP cyclohydrolase II / DHBP synthase activity, wherein said DNA molecule comprises a 20 base pair nucleotide portion identical in sequence to a consecutive 20 base pair portion of the coding sequence set forth in SEQ ID NO:13. In still another example, the DNA molecule of the invention comprises the coding sequence set forth in SEQ ID NO:13 and encodes an enzyme having bifunctional GTP cyclohydrolase II / DHBP synthase activity. Although the nucleotide sequence provided in SEQ ID NO:13 that encodes the bifunctional GTP cyclohydrolase II / DHBP synthase was isolated from *Arabidopsis thaliana*, using the information provided by the present invention, the nucleotide sequence that encodes an enzyme having bifunctional GTP cyclohydrolase II / DHBP synthase activity can be obtained from any plant using standard methods known in the art.

The present invention also provides a chimeric gene comprising a promoter operatively linked to a DNA molecule of the invention. Further, the present invention provides a recombinant vector comprising such a chimeric gene, wherein the vector is capable of being stably transformed into a host cell. Still further, the present invention provides a host cell comprising such a vector, wherein the host cell is capable of expressing the DNA molecule encoding an enzyme involved in riboflavin biosynthesis. A host cell according to the invention may be a bacterial cell, a yeast cell, or a plant cell. Especially the host cell according to the invention may be a bacterial cell.

The present invention further provides a process for producing nucleotides sequences encoding gene products having altered lumazine synthase activity comprising: (a) shuffling a DNA molecule from a plant that encodes an enzyme having lumazine synthase acitivity, (b) expressing the resulting shuffled nucleotide sequences, and (c) selecting for altered lumazine synthase activity as compared to the activity of an enzyme encoded by the unshuffled DNA molecule. Preferably, the nucleotide sequence shuffled according to this method is SEQ ID NO: 1. The invention is also directed to a shuffled DNA molecule obtainable by this process. Preferably, the shuffled DNA molecule encodes an enzyme having enhanced tolerance to an inhibitor of lumazine synthase activity. The present invention also provides a chimeric gene comprising a promoter operatively linked to a shuffled DNA molecule; a recombinant vector comprising said chimeric gene, wherein said vector is capable of being stably transformed into a host cell; a host cell comprising said vector. Said host cell is preferably a bacterial cell, a yeast cell, or a plant cell, especially a plant cell. The invention is also directed to a plant or seed comprising such a plant cell. Preferably, said plant is tolerant to an inhibitor of lumazine synthase activity.

The present invention further provides a process for producing nucleotides sequences encoding gene products having altered bifunctional GTP cyclohydrolase II / DHBP synthase activity comprising: (a) shuffling a DNA molecule from a plant that encodes an enzyme having bifunctional GTP cyclohydrolase II / DHBP synthase acitivity, (b) expressing the resulting shuffled nucleotide sequences, and (c) selecting for altered bifunctional GTP cyclohydrolase II / DHBP synthase activity as compared to the activity of an enzyme encoded by the unshuffled DNA molecule. Preferably, the nucleotide sequence shuffled according to this method is SEQ ID NO: 13. The invention is also directed to a shuffled DNA molecule obtainable by this process. Preferably, the shuffled DNA molecule encodes an enzyme having enhanced tolerance to an inhibitor of bifunctional GTP cyclohydrolase II / DHBP synthase activity. The present invention also provides a chimeric gene comprising a promoter operatively linked to a shuffled DNA molecule; a recombinant vector comprising said chimeric gene, wherein said vector is capable of being stably transformed into a host cell; a host cell comprising said vector. Said host cell is preferably a bacterial cell, a yeast cell, or a plant cell, especially a plant cell. The invention is also directed to a plant or seed comprising such a plant cell. Preferably, said plant is tolerant to an inhibitor of bifunctional GTP cyclohydrolase II / DHBP synthase activity.

In accordance with another embodiment, the present invention also relates to the recombinant production of the above-described riboflavin biosynthesis enzymes and

methods of use thereof. In particular, the present invention provides an isolated plant enzyme involved in riboflavin biosynthesis, wherein the enzyme has lumazine synthase activity. Preferably this enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:2. More preferably, this enzyme comprises the amino acid sequence set forth in SEQ ID NO:2. The present invention also provides an isolated plant enzyme involved in riboflavin biosynthesis, wherein the enzyme has bifunctional GTP cyclohydrolase II / DHBP synthase activity. Preferably, this enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:14. More preferably, this enzyme comprises the amino acid sequence set forth in SEQ ID NO:14.

The present invention further provides methods of using purified plant riboflavin biosynthesis enzymes such as lumazine synthase and bifunctional GTP cyclohydrolase II / DHBP synthase to screen for novel inhibitors thereof, which can then be used as herbicides to suppress the growth of undesirable vegetation in fields where crops are grown, particularly agronomically important crops such as maize and other cereal crops such as wheat, oats, rye, sorghum, rice, barley, millet, turf and forage grasses, and the like, as well as cotton, sugar cane, sugar beet, oilseed rape, and soybeans.

With regard to lumazine synthase, such a screen for chemicals having the ability to inhibit lumazine synthase activity preferably comprises the steps of: (a) combining an enzyme having lumazine synthase activity in a first reaction mixture with 2,4-dioxy-5-amino-6-ribitylamino-pyrimidine and 3,4-dihydroxy-2-butanone phosphate under conditions in which the enzyme is capable of catalyzing the synthesis of lumazine; (b) combining the chemical and the enzyme in a second reaction mixture with 2,4-dioxy-5-amino-6-ribitylamino-pyrimidine and 3,4-dihydroxy-2-butanone phosphate under the same conditions as in the first reaction mixture; (c) determining the amounts of lumazine produced in the first and second reaction mixtures; and (d) comparing the amounts of lumazine produced in the first and second reaction mixtures; wherein the chemical is capable of inhibiting the lumazine synthase activity of the enzyme if the amount of lumazine produced in the second reaction mixture is significantly less than the amount of lumazine produced in the first reaction mixture. Preferred is a method for screening according to the invention wherein the first reaction mixture comprises 50μM 2,4-dioxy-5-amino-6-ribitylamino-pyrimidine, and 0.5 mM 3,4-dihydroxy-2-butanone phosphate. Further preferred is a method for screening according to the invention, wherein the amounts of lumazine produced in the reaction mixtures are determined using a fluorimeter at an excitation wavelength of 407 nm.

With regard to the bifunctional GTP cyclohydrolase II / DHBP synthase, such a screen for chemicals having the ability to inhibit GTP cyclohydrolase II / DHBP synthase synthase activity preferably comprises the steps of: (a) combining an enzyme having GTP cyclohydrolase II / DHBP synthase activity in a first reaction mixture with GTP or ribulose-5-phosphate under conditions in which the enzyme is capable of catalyzing the synthesis of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate, respectively; (b) combining the chemical and the enzyme in a second reaction mixture with GTP or ribulose-5-phosphate under the same conditions as in the first reaction mixture; (c) determining the amounts of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate produced in the first and second reaction mixtures; and (d) comparing the amounts of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate produced in the first and second reaction mixtures; wherein the chemical is capable of inhibiting the bifunctional GTP cyclohydrolase II / DHBP synthase activity of the enzyme if the amount of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate produced in the second reaction mixture is significantly less than the amount of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate produced in the first reaction mixture.

The present invention also embodies herbicidal chemicals identified by the above screening methods in addition to methods for suppressing the growth of plants by applying such herbicidal chemicals to the plants, whereby the chemicals inhibit the activity of lumazine synthase or bifunctional GTP cyclohydrolase II / DHBP synthase in the plants.

The present invention further embodies plants, plant tissues, plant seeds, and plant cells that have modified riboflavin biosynthesis enzyme activity and that are therefore tolerant to inhibition by a herbicide at levels normally inhibitory to naturally occurring riboflavin biosynthesis enzyme activity. Herbicide tolerant plants encompassed by the invention include those that would otherwise be potential targets for normally inhibiting herbicides, particularly the agronomically important crops mentioned above. According to this embodiment, plants, plant tissue, plant seeds, or plant cells are stably transformed with a recombinant DNA molecule comprising a suitable promoter functional in plants operatively linked to a nucleotide coding sequence that encodes a modified riboflavin biosynthesis enzyme that is tolerant to inhibition by a herbicide at a concentration that would normally inhibit the activity of wild-type, unmodified riboflavin biosynthesis enzyme. Modified

riboflavin biosynthesis enzyme activity may also be conferred upon a plant by increasing expression of wild-type herbicide-sensitive riboflavin biosynthesis enzyme by providing multiple copies of wild-type riboflavin biosynthesis genes to the plant or by overexpression of wild-type riboflavin biosynthesis genes under control of a stronger-than-wild-type promoter. The transgenic plants, plant tissue, plant seeds, or plant cells thus created are then selected by conventional selection techniques, whereby herbicide tolerant lines are isolated, characterized, and developed. Alternately, random or site-specific mutagenesis may be used to generate herbicide tolerant lines.

Therefore, the present invention provides a plant, plant cell, plant seed, or plant tissue comprising a DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme involved in riboflavin biosynthesis, wherein the enzyme has lumazine synthase activity and wherein the DNA molecule confers upon the plant, plant cell, plant seed, or plant tissue tolerance to a herbicide in amounts that normally naturally occurring lumazine synthase activity. According to one example of this embodiment, the enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:2. According to another example of this embodiment, the DNA molecule is substantially similar to the coding sequence set forth in SEQ ID NO:1. In a related aspect, the present invention is directed to a method for selectively suppressing the growth of weeds in a field containing a crop of planted crop seeds or plants, comprising the steps of: (a) planting herbicide tolerant crops or crop seeds, which are plants or plant seeds that are tolerant to a herbicide that inhibits naturally occurring lumazine synthase activity; and (b) applying to the crops or crop seeds and the weeds in the field a herbicide in amounts that inhibit naturally occurring lumazine synthase activity, wherein the herbicide suppresses the growth of the weeds without significantly suppressing the growth of the crops.

The present invention further provides a plant, plant cell, plant seed, or plant tissue comprising a DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme involved in riboflavin biosynthesis, wherein the enzyme has bifunctional GTP cyclohydrolase II / DHBP synthase activity and wherein the DNA molecule confers upon the plant, plant cell, plant seed, or plant tissue tolerance to a herbicide in amounts that normally naturally occurring bifunctional GTP cyclohydrolase II / DHBP synthase activity. According to one example of this embodiment, the enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:14. According to another example of this embodiment, the DNA molecule is

substantially similar to the coding sequence set forth in SEQ ID NO:13. In a related aspect, the present invention is directed to a method for selectively suppressing the growth of weeds in a field containing a crop of planted crop seeds or plants, comprising the steps of: (a) planting herbicide tolerant crops or crop seeds, which are plants or plant seeds that are tolerant to a herbicide that inhibits naturally occurring bifunctional GTP cyclohydrolase II / DHBP synthase activity; and (b) applying to the crops or crop seeds and the weeds in the field a herbicide in amounts that inhibit naturally occurring bifunctional GTP cyclohydrolase II / DHBP synthase activity, wherein the herbicide suppresses the growth of the weeds without significantly suppressing the growth of the crops.

Other objects and advantages of the present invention will become apparent to those skilled in the art from a study of the following description of the invention and non-limiting examples.

I. Plant Riboflavin Biosynthesis Genes

In one aspect, the present invention is directed to a DNA molecule comprising a nucleotide sequence isolated from a plant source that encodes the β subunit of riboflavin synthase (lumazine synthase). In particular, the present invention provides a DNA molecule isolated from *Arabidopsis thaliana* that encodes lumazine synthase and DNA molecules substantially similar thereto that encode enzymes having lumazine synthase activity. The DNA coding sequence for lumazine synthase from *Arabidopsis thaliana* is provided in SEQ ID NO:1.

In another aspect, the present invention is directed to a DNA molecule comprising a nucleotide sequence isolated from a plant source that encodes the bifunctional enzyme GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone phosphate (DHBP). In particular, the present invention provides a DNA molecule isolated from *Arabidopsis thaliana* that encodes this bifunctional enzyme and DNA molecules substantially similar thereto that encode enzymes having GTP cyclohydrolase II / DHBP synthase activity. The DNA coding sequence for GTP cyclohydrolase II / DHBP synthase from *Arabidopsis thaliana* is provided in SEQ ID NO:13. The present invention represents the first recognition that in plants, GTP cyclohydrolase II and DHBP synthase constitute a single, bifunctional enzyme

Based on Applicants' disclosure of the present invention, DNA sequences encoding riboflavin biosynthesis enzymes can, for the first time, be isolated from the genome of any desired plant species. An exemplary method for isolating riboflavin biosynthesis genes from

plants is described in Examples 1 and 11. With this method, searches of the *Arabidopsis thaliana* Expressed Sequence Tag (EST) database (Arabidopsis Biological Resource Center at Ohio State, Ohio State University, Columbus, OH) revealed ESTs with homologies to the *E.coli* riboflavin synthase β subunit and *B. subtilis* GTP cyclohydrolase. DNA fragments generated by PCR with primers specific to these ESTs were used to probe an *Arabidopsis* lambda ZAP library, whereupon cDNAs were isolated. The determined protein sequence encoded by one cDNA showed approximately 68% similarity to both the *E. coli* and *B. subtilis* riboflavin synthase β subunit. The determined protein sequence encoded by another cDNA showed approximately 70% similarity to the *B. subtilis* GTP cyclohydrolase.

Alternatively, riboflavin biosynthesis gene sequences can be isolated from any plant according to well known techniques based on their sequence similarity to the *Arabidopsis thaliana* coding sequences (SEQ ID NOs:1 and 13) taught by the present invention. In these techniques, all or part of a known plant riboflavin biosynthesis gene's coding sequence is used as a probe that selectively hybridizes to other riboflavin biosynthesis gene sequences present in a population of cloned genomic DNA fragments or cDNA fragments (i.e. genomic or cDNA libraries) from a chosen plant. Such techniques include hybridization screening of plated DNA libraries (either plaques or colonies; see, *e.g.*. Sambrook *et al.*, "Molecular Cloning", eds., Cold Spring Harbor Laboratory Press. (1989)) and amplification by PCR using oligonucleotide primers corresponding to sequence domains conserved among known riboflavin biosynthesis enzyme's amino acid sequences (see, *e.g.* Innis *et al.*, "PCR Protocols, a Guide to Methods and Applications", pub. by Academic Press (1990)). These methods are particularly well suited to the isolation of riboflavin biosynthesis gene sequences from organisms closely related to the organism from which the probe sequence is derived. Thus, application of these methods using the *Arabidopsis* coding sequences as probes would be expected to be particularly well suited for the isolation of riboflavin biosynthesis gene sequences from other plant species, including monocotyledons and dicotyledons.

The isolated riboflavin biosynthesis gene sequences taught by the present invention can be manipulated according to standard genetic engineering techniques to suit any desired purpose. For example, an entire plant riboflavin biosynthesis gene sequence or portions thereof may be used as a probe capable of specifically hybridizing to coding sequences and messenger RNAs. To achieve specific hybridization under a variety of

conditions, such probes include sequences that are unique among plant riboflavin biosynthesis gene sequences and are at least 10 nucleotides in length, preferably at least 20 nucleotides in length, and most preferably at least 50 nucleotides in length. Such probes may be used to amplify and analyze riboflavin biosynthesis gene sequences from a chosen organism via PCR. This technique may be useful to isolate additional riboflavin biosynthesis gene sequences from a desired organism or as a diagnostic assay to determine the presence of riboflavin biosynthesis gene sequences in an organism. This technique may also be used to detect the presence of altered riboflavin biosynthesis gene sequences associated with a particular condition of interest such as herbicide tolerance, poor health, etc.

Lumazine synthase-specific and GTP cyclohydrolase II / DHBP synthase-specific hybridization probes can also be used to map the location of these native genes in the genome of a chosen plant using standard techniques based on the selective hybridization of the probe to genomic sequences. These techniques include, but are not limited to, identification of DNA polymorphisms identified or contained within the probe sequence, and use of such polymorphisms to follow segregation of the gene relative to other markers of known map position in a mapping population derived from self fertilization of a hybrid of two polymorphic parental lines (see e.g. Helentjaris *et al.*, *Plant Mol. Biol. 5:* 109 (1985); Sommer *et al. Biotechniques 12:*82 (1992); D'Ovidio *et al.*, *Plant Mol. Biol. 15:* 169 (1990)). While any plant riboflavin biosynthesis gene sequence is contemplated to be useful as a probe for mapping riboflavin biosynthesis genes, preferred probes are those gene sequences from plant species more closely related to the chosen plant species, and most preferred probes are those gene sequences from the chosen plant species. Mapping of riboflavin biosynthesis genes in this manner is contemplated to be particularly useful for breeding purposes. For instance, by knowing the genetic map position of a mutant riboflavin biosynthesis gene that confers herbicide resistance, flanking DNA markers can be identified from a reference genetic map (see, *e.g.*, Helentjaris, *Trends Genet. 3:* 217 (1987)). During introgression of the herbicide resistance trait into a new breeding line, these markers can then be used to monitor the extent of linked flanking chromosomal DNA still present in the recurrent parent after each round of back-crossing.

Lumazine synthase-specific and GTP cyclohydrolase II / DHBP synthase-specific hybridization probes can also be used to quantify levels of riboflavin biosynthesis gene mRNA in a plant using standard techniques such as Northern blot analysis. This technique is useful as a diagnostic assay to detect altered levels of riboflavin biosynthesis gene

expression that are associated with particular conditions such as enhanced tolerance to herbicides that target riboflavin biosynthesis genes.

II. Essentiality of Riboflavin Biosynthesis Genes in Plants Demonstrated by Antisense Inhibition

As shown in the examples below, the essentiality of riboflavin biosynthesis genes for normal plant growth and development has been demonstrated by antisense inhibition of lumazine synthase in plants using the antisense validation system described in co-owned and co-pending application serial no. 08/978,830 [entitled "Methods and Compositions Useful for the Activation of Silent Transgenes", filed Nov. 26, 1997], incorporated herein by reference. In this system, a hybrid transcription factor gene is made that comprises a DNA-binding domain and an activation domain. In addition, an activatable DNA construct is made that comprises a synthetic promoter operatively linked to an activatable DNA sequence. The hybrid transcription factor gene and synthetic promoter are selected or designed such that the DNA binding domain of the hybrid transcription factor is capable of binding specifically to the synthetic promoter, which then activates expression of the activatable DNA sequence. A first plant is transformed with the hybrid transcription factor gene, and a second plant is transformed with the activatable DNA construct. The first plant and second plants are crossed to produce a progeny plant containing both the sequence encoding the hybrid transcription factor and the synthetic promoter, wherein the activatable DNA sequence is expressed in the progeny plant. In the preferred embodiment, the activatable DNA sequence is an antisense sequence capable of inactivating expression of an endogenous gene such as the lumazine synthase gene or the bifunctional GTP cyclohydrolase II / DHBP synthase gene. Hence, the progeny plant will be unable to normally express the endogenous gene.

This antisense validation system is especially useful for allowing expression of traits that might otherwise be unrecoverable as constitutively driven transgenes. For instance, foreign genes with potentially lethal effect or antisense genes or dominant-negative mutations designed to abolish function of essential genes, while of great interest in basic studies of plant biology, present inherent experimental problems. Decreased transformation frequencies are often cited as evidence of lethality associated with a particular constitutively driven transgene, but negative results of this type are laden with alternative trivial explanations. The present invention is an important advancement in the field of agriculture because it allows stable maintenance and propagation of a test transgene separate from its

expression. This ability to separate transgene insertion from expression is especially useful for firm conclusions about essentiality of gene function to be drawn. A substantial benefit of the present invention is that plant genes essential for normal growth or development can thus be identified in this manner. The identification of such genes provide useful targets for screening compound libraries for effective herbicides. Below, the antisense validation system is described in greater detail:

A. Hybrid Transcription Factor Gene

A hybrid transcription factor gene for use in the antisense validation system described herein comprises DNA sequences encoding (1) a DNA-binding domain and (2) an activation domain that interacts with components of transcriptional machinery assembling at a promoter. Gene fragments are joined, typically such that the DNA binding domain is toward the 5' terminus and the activator domain is toward the 3' terminus, to form a hybrid gene whose expression produces a hybrid transcription factor. One skilled in the art is capable of routinely combining various DNA sequences encoding DNA binding domains with various DNA sequences encoding activation domains to produce a wide array of hybrid transcription factor genes. Examples of DNA sequences encoding DNA binding domains include, but are not limited to, those encoding the DNA binding domains of GAL4, bacteriophage 434, *lexA*, *lacI*, and phage lambda repressor. Examples of DNA sequences encoding the activation domain include, but are not limited to, those encoding the acidic activation domains of herpes simplex VP16, maize C1, and P1. In addition, suitable activation domains can be isolated by fusing DNA pieces from an organism of choice to a suitable DNA binding domain and selecting directly for function (Estruch *et al.*, (1994) *Nucleic Acids Res.* 22: 3983-3989). Domains of transcriptional activator proteins can be swapped between proteins of diverse origin (Brent and Ptashne (1985) *Cell* 43: 729-736). A preferable hybrid transcription factor gene comprises DNA sequences encoding the GAL4 DNA binding domain fused to the maize C1 activation domain.

B. Activatable DNA Construct

An activatable DNA construct for use in the antisense validation system described herein comprises (1) a synthetic promoter operatively linked to (2) an activatable DNA sequence. The synthetic promoter comprises at least one DNA binding site recognized by the DNA binding domain of the hybrid transcription factor, and a minimal promoter, preferably a TATA element derived from a promoter recognized by plant cells. More

promoter, signal sequence, 5' and 3' untranslated sequences, and enhancer appropriate for the chosen host is within the level of skill of the routineer in the art. The resultant molecule, containing the individual elements linked in proper reading frame, may be inserted into a vector capable of being transformed into the host cell. Suitable expression vectors and methods for recombinant production of proteins are well known for host organisms such as *E. coli*, yeast, and insect cells (see, *e.g.*, Luckow and Summers, *Bio/Technol. 6:* 47 (1988)). Specific examples include plasmids such as pBluescript (Stratagene, La Jolla, CA), pFLAG (International Biotechnologies, Inc., New Haven, CT), pTrcHis (Invitrogen, La Jolla, CA), and baculovirus expression vectors, e.g., those derived from the genome of *Autographica californica* nuclear polyhedrosis virus (AcMNPV). A preferred baculovirus/insect system is pVl11392/Sf21 cells (Invitrogen, La Jolla, CA).

Recombinantly produced plant riboflavin biosynthesis enzymes can be isolated and purified using a variety of standard techniques. The actual techniques that may be used will vary depending upon the host organism used, whether the enzyme is designed for secretion, and other such factors familiar to the skilled artisan (*see, e.g.* chapter 16 of Ausubel, F. *et al.*, "Current Protocols in Molecular Biology", pub. by John Wiley & Sons, Inc. (1994).

Recombinantly produced plant riboflavin biosynthesis enzymes are useful for a variety of purposes. For example, they can be used in *in vitro* assays to screen known herbicidal chemicals whose target has not been identified to determine if they inhibit riboflavin biosynthesis enzymes. Such *in vitro* assays may also be used as more general screens to identify chemicals that inhibit such enzymatic activity and that are therefore herbicide candidates. Alternatively, recombinantly produced riboflavin biosynthesis enzymes may be used to further characterize their association with known inhibitors in order to rationally design new inhibitory herbicides as well as herbicide tolerant forms of the enzymes.

Inhibitor Assay:

Thus, an assay useful for identifying inhibitors of essential plant genes, such as plant riboflavin biosynthesis genes, comprises the steps of:

a) reacting a plant riboflavin biosynthesis enzyme and a substrate thereof in the presence of a suspected inhibitor of the enzyme's function;

b) comparing the rate of enzymatic activity in the presence of the suspected inhibitor to the rate of enzymatic activity under the same conditions in the absence of the suspected inhibitor; and

c) determining whether the suspected inhibitor inhibits the riboflavin biosynthesis enzyme.

For example, the inhibitory effect on plant lumazine synthase may determined by a reduction or complete inhibition of lumazine synthesis in the assay. Such a determination may be made by comparing, in the presence and absence of the candidate inhibitor, the amount of lumazine synthesized in the *in vitro* assay using fluorescence or absorbance detection as described *infra* in the Examples. A similar assay may be used to screen for inhibitors of the bifunctional plant GTP cyclohydrolase II / DHBP synthase enzyme.

In addition, recombinantly produced plant riboflavin biosynthesis enzymes may be used to elucidate the complex structure of these molecules, such as has been done for riboflavin synthase from *Bacillus subtilis* (Ladenstein, *et al.*, (1988) *J. Mol. Biol.* 203, 1045-1070). Such information regarding the structure of the plant riboflavin biosynthesis enzymes may be used, for example, in the rational design of new inhibitory herbicides.

IV: Herbicide Tolerant Plants

The present invention is further directed to plants, plant tissue, plant seeds, and plant cells tolerant to herbicides that inhibit the naturally occurring riboflavin biosynthesis in these plants, wherein the tolerance is conferred by altered riboflavin biosynthesis enzyme activity. Altered riboflavin biosynthesis enzyme activity may be conferred upon a plant according to the invention by increasing expression of wild-type herbicide-sensitive riboflavin biosynthesis enzyme by providing additional wild-type riboflavin biosynthesis genes to the plant, by expressing modified herbicide-tolerant riboflavin biosynthesis enzymes in the plant, or by a combination of these techniques. Representative plants include any plants to which these herbicides are applied for their normally intended purpose. Preferred are agronomically important crops such as cotton, soybean, oilseed rape, sugar beet, maize, rice, wheat, barley, oats, rye, sorghum, millet, turf, forage, turf grasses, and the like.

A. Increased Expression of Wild-Type Riboflavin Biosynthesis Enzymes

Achieving altered riboflavin biosynthesis enzyme activity through increased expression results in a level of a riboflavin biosynthesis enzyme in the plant cell at least sufficient to overcome growth inhibition caused by the herbicide. The level of expressed

enzyme generally is at least two times, preferably at least five times, and more preferably at least ten times the natively expressed amount. Increased expression may be due to multiple copies of a wild-type riboflavin biosynthesis gene; multiple occurrences of the coding sequence within the gene (*i.e.* gene amplification) or a mutation in the non-coding, regulatory sequence of the endogenous gene in the plant cell. Plants having such altered gene activity can be obtained by direct selection in plants by methods known in the art (see, *e.g.* U.S. Patent No. 5,162,602, and U.S. Patent No. 4,761,373, and references cited therein). These plants also may be obtained by genetic engineering techniques known in the art. Increased expression of a herbicide-sensitive riboflavin biosynthesis gene can also be accomplished by stably transforming a plant cell with a recombinant or chimeric DNA molecule comprising a promoter capable of driving expression of an associated structural gene in a plant cell operatively linked to a homologous or heterologous structural gene encoding the riboflavin biosynthesis enzyme.

B. Expression of Modified Herbicide-Tolerant Riboflavin Biosynthesis Enzymes

According to this embodiment, plants, plant tissue, plant seeds, or plant cells are stably transformed with a recombinant DNA molecule comprising a suitable promoter functional in plants operatively linked to a coding sequence encoding a herbicide tolerant form of a riboflavin biosynthesis enzyme. A herbicide tolerant form of the enzyme has at least one amino acid substitution, addition or deletion that confers tolerance to a herbicide that inhibits the unmodified, naturally occurring form of the enzyme. The transgenic plants, plant tissue, plant seeds, or plant cells thus created are then selected by conventional selection techniques, whereby herbicide tolerant lines are isolated, characterized, and developed. Below are described methods for obtaining genes that encode herbicide tolerant forms of riboflavin biosynthesis enzymes:

One general strategy involves direct or indirect mutagenesis procedures on microbes. For instance, a genetically manipulatable microbe such as *E. coli* or *S. cerevisiae* may be subjected to random mutagenesis *in vivo* with mutagens such as UV light or ethyl or methyl methane sulfonate. Mutagenesis procedures are described, for example, in Miller, *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1972); Davis *et al.*, *Advanced Bacterial Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1980); Sherman *et al.*, *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1983); and U.S. Patent No. 4,975,374. The microbe

selected for mutagenesis contains a normal, inhibitor-sensitive riboflavin biosynthesis gene and is dependent upon the activity conferred by this gene. The mutagenized cells are grown in the presence of the inhibitor at concentrations that inhibit the unmodified gene. Colonies of the mutagenized microbe that grow better than the unmutagenized microbe in the presence of the inhibitor (i.e. exhibit resistance to the inhibitor) are selected for further analysis. Riboflavin biosynthesis genes from these colonies are isolated, either by cloning or by PCR amplification, and their sequences are elucidated. Sequences encoding altered gene products are then cloned back into the microbe to confirm their ability to confer inhibitor tolerance.

A method of obtaining mutant herbicide-tolerant alleles of a plant riboflavin biosynthesis gene involves direct selection in plants. For example, the effect of a mutagenized riboflavin biosynthesis gene on the growth inhibition of plants such as *Arabidopsis*, soybean, or maize is determined by plating seeds sterilized by art-recognized methods on plates on a simple minimal salts medium containing increasing concentrations of the inhibitor. Such concentrations are in the range of 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 110, 300, 1000 and 3000 parts per million (ppm). The lowest dose at which significant growth inhibition can be reproducibly detected is used for subsequent experiments.

Mutagenesis of plant material is utilized to increase the frequency at which resistant alleles occur in the selected population. Mutagenized seed material is derived from a variety of sources, including chemical or physical mutagenesis or seeds, or chemical or physical mutagenesis or pollen (Neuffer, In *Maize for Biological Research* Sheridan, ed. Univ. Press, Grand Forks, ND., pp. 61-64 (1982)), which is then used to fertilize plants and the resulting M_1 mutant seeds collected. Typically for *Arabidopsis*, M_2 seeds (Lehle Seeds, Tucson, AZ), which are progeny seeds of plants grown from seeds mutagenized with chemicals, such as ethyl methane sulfonate, or with physical agents, such as gamma rays or fast neutrons, are plated at densities of up to 10,000 seeds/plate (10 cm diameter) on minimal salts medium containing an appropriate concentration of inhibitor to select for tolerance. Seedlings that continue to grow and remain green 7-21 days after plating are transplanted to soil and grown to maturity and seed set. Progeny of these seeds are tested for tolerance to the herbicide. If the tolerance trait is dominant, plants whose seed segregate 3:1 / resistant:sensitive are presumed to have been heterozygous for the resistance at the M_2 generation. Plants that give rise to all resistant seed are presumed to have been homozygous for the resistance at the M_2 generation. Such mutagenesis on

intact seeds and screening of their M2 progeny seed can also be carried out on other species, for instance soybean (*see, e.g.* U.S. Pat. No. 5,084,082). Alternatively, mutant seeds to be screened for herbicide tolerance are obtained as a result of fertilization with pollen mutagenized by chemical or physical means.

Confirmation that the genetic basis of the herbicide tolerance is a modified riboflavin biosynthesis gene is ascertained as exemplified below. First, alleles of the riboflavin biosynthesis gene from plants exhibiting resistance to the inhibitor are isolated using PCR with primers based either upon conserved regions in the *Arabidopsis* cDNA coding sequences shown in SEQ ID NO:1 or SEQ ID NO:13 or, more preferably, based upon the unaltered riboflavin biosynthesis gene sequence from the plant used to generate tolerant alleles. After sequencing the alleles to determine the presence of mutations in the coding sequence, the alleles are tested for their ability to confer tolerance to the inhibitor on plants into which the putative tolerance-conferring alleles have been transformed. These plants can be either *Arabidopsis* plants or any other plant whose growth is susceptible to the inhibitors. Second, the riboflavin biosynthesis genes are mapped relative to known restriction fragment length polymorphisms (RFLPs) (*See, for example*, Chang *et al. Proc. Natl. Acad, Sci, USA* 85: 6856-6860 (1988); Nam *et al., Plant Cell* 1: 699-705 (1989). The tolerance trait is independently mapped using the same markers. When tolerance is due to a mutation in that riboflavin biosynthesis gene, the tolerance trait maps to a position indistinguishable from the position of the riboflavin biosynthesis gene.

Another method of obtaining herbicide-tolerant alleles of a riboflavin biosynthesis gene is by selection in plant cell cultures. Explants of plant tissue, *e.g.* embryos, leaf disks, etc. or actively growing callus or suspension cultures of a plant of interest are grown on medium in the presence of increasing concentrations of the inhibitory herbicide or an analogous inhibitor suitable for use in a laboratory environment. Varying degrees of growth are recorded in different cultures. In certain cultures, fast-growing variant colonies arise that continue to grow even in the presence of normally inhibitory concentrations of inhibitor. The frequency with which such faster-growing variants occur can be increased by treatment with a chemical or physical mutagen before exposing the tissues or cells to the inhibitor. Putative tolerance-conferring alleles of the riboflavin biosynthesis gene are isolated and tested as described in the foregoing paragraphs. Those alleles identified as conferring herbicide tolerance may then be engineered for optimal expression and transformed into

the plant. Alternatively, plants can be regenerated from the tissue or cell cultures containing these alleles.

Still another method involves mutagenesis of wild-type, herbicide sensitive plant riboflavin biosynthesis genes in bacteria or yeast, followed by culturing the microbe on medium that contains inhibitory concentrations of the inhibitor and then selecting those colonies that grow in the presence of the inhibitor. More specifically, a plant cDNA, such as the *Arabidopsis* cDNA encoding lumazine synthase (SEQ ID NO:1) or the bifunctional GTP cyclohydrolase II / DHBP synthase enzyme (SEQ ID NO:13) is cloned into a microbe that otherwise lacks the selected gene's activity. The transformed microbe is then subjected to *in vivo* mutagenesis or to *in vitro* mutagenesis by any of several chemical or enzymatic methods known in the art, e.g. sodium bisulfite (Shortle *et al., Methods Enzymol. 100*:457-468 (1983); methoxylamine (Kadonaga *et al., Nucleic Acids Res. 13*:1733-1745 (1985); oligonucleotide-directed saturation mutagenesis (Hutchinson *et al., Proc. Natl. Acad. Sci. USA, 83*:710-714 (1986); or various polymerase misincorporation strategies (see, e.g. Shortle et al., Proc. Natl. Acad. Sci. USA, 79:1588-1592 (1982); Shiraishi *et al., Gene 64*:313-319 (1988); and Leung *et al., Technique 1*:11-15 (1989). Colonies that grow in the presence of normally inhibitory concentrations of inhibitor are picked and purified by repeated restreaking. Their plasmids are purified and tested for the ability to confer tolerance to the inhibitor by retransforming them into the microbe lacking riboflavin biosynthesis gene activity. The DNA sequences of cDNA inserts from plasmids that pass this test are then determined.

Herbicide resistant riboflavin biosynthesis genes are also obtained using methods involving *in vitro* recombination, also called DNA shuffling. By DNA shuffling, mutations, preferably random mutations, are introduced in riboflavin biosynthesis genes. DNA shuffling also leads to the recombination and rearrangement of sequences within a riboflavin biosynthesis gene or to recombination and exchange of sequences between two or more different riboflavin biosynthesis protein encoding sequences. These methods allow for the production of millions of mutated riboflavin biosynthesis genes. The mutated genes, or shuffled genes, are screened for desirable properties, e.g. improved tolerance to herbicides and for mutations that provide broad spectrum tolerance to the different classes of inhibitor chemistry. Such screens are well within the skills of a routineer in the art.

In a preferred embodiment, a mutagenized riboflavin biosynthesis gene is formed from at least one template riboflavin biosynthesis gene, wherein the template riboflavin

biosynthesis gene has been cleaved into double-stranded random fragments of a desired size, and comprising the steps of adding to the resultant population of double-stranded random fragments one or more single or double-stranded oligonucleotides, wherein said oligonucleotides comprise an area of identity and an area of heterology to the double-stranded random fragments; denaturing the resultant mixture of double-stranded random fragments and oligonucleotides into single-stranded fragments; incubating the resultant population of single-stranded fragments with a polymerase under conditions which result in the annealing of said single-stranded fragments at said areas of identity to form pairs of annealed fragments, said areas of identity being sufficient for one member of a pair to prime replication of the other, thereby forming a mutagenized double-stranded polynucleotide; and repeating the second and third steps for at least two further cycles, wherein the resultant mixture in the second step of a further cycle includes the mutagenized double-stranded polynucleotide from the third step of the previous cycle, and the further cycle forms a further mutagenized double-stranded polynucleotide, wherein the mutagenized polynucleotide is a mutated riboflavin biosynthesis gene having enhanced tolerance to a herbicide which inhibits naturally occurring riboflavin biosynthesis activity. In a preferred embodiment, the concentration of a single species of double-stranded random fragment in the population of double-stranded random fragments is less than 1% by weight of the total DNA. In a further preferred embodiment, the template double-stranded polynucleotide comprises at least about 100 species of polynucleotides. In another preferred embodiment, the size of the double-stranded random fragments is from about 5 bp to 5 kb. In a further preferred embodiment, the fourth step of the method comprises repeating the second and the third steps for at least 10 cycles. Such method is described e.g. in Stemmer et al. (1994) Nature 370: 389-391, in US Patent 5,605,793 and in Crameri et al. (1998) Nature 391: 288-291, as well as in WO 97/20078, and these references are incorporated herein by reference.

In another preferred embodiment, any combination of two or more different riboflavin biosynthesis genes are mutagenized *in vitro* by a staggered extension process (StEP), as described e.g. in Zhao et al. (1998) Nature Biotechnology 16: 258-261. Briefly, the two or more riboflavin biosynthesis genes are used as template for PCR amplification with the extension cycles of the PCR reaction preferably carried out at a lower temperature than the optimal polymerization temperature of the polymerase. For example, when a thermostable polymerase with an optimal temperature of approximately 72°C is used, the temperature for the extension reaction is desirably below 72°C, more desirably below 65°C, preferably

below 60°C, more preferably the temperature for the extension reaction is 55°C. Additionally, the duration of the extension reaction of the PCR cycles is desirably shorter than usually carried out in the art, more desirably it is less than 30 seconds, preferably it is less than 15 seconds, more preferably the duration of the extension reaction is 5 seconds. Only a short DNA fragment is polymerized in each extension reaction, allowing template switch of the extension products between the starting DNA molecules after each cycle of denaturation and annealing, thereby generating diversity among the extension products. The optimal number of cycles in the PCR reaction depends on the length of the riboflavin biosynthesis coding regions to be mutagenized but desirably over 40 cycles, more desirably over 60 cycles, preferably over 80 cycles are used. Optimal extension conditions and the optimal number of PCR cycles for every combination of riboflavin biosynthesis genes are determined as described in using procedures well-known in the art. The other parameters for the PCR reaction are essentially the same as commonly used in the art. The primers for the amplification reaction are preferably designed to anneal to DNA sequences located outside of the coding sequence of the riboflavin biosynthesis genes, e.g. to DNA sequences of a vector comprising the riboflavin biosynthesis genes, whereby the different riboflavin biosynthesis genes used in the PCR reaction are preferably comprised in separate vectors. The primers desirably anneal to sequences located less than 500 bp away from riboflavin biosynthesis coding sequences, preferably less than 200 bp away from the riboflavin biosynthesis coding sequences, more preferably less than 120 bp away from the riboflavin biosynthesis coding sequences. Preferably, the riboflavin biosynthesis coding sequences are surrounded by restriction sites, which are included in the DNA sequence amplified during the PCR reaction, thereby facilitating the cloning of the amplified products into a suitable vector.

In another preferred embodiment, fragments of riboflavin biosynthesis genes having cohesive ends are produced as described in WO 98/05765. The cohesive ends are produced by ligating a first oligonucleotide corresponding to a part of a riboflavin biosynthesis gene to a second oligonucleotide not present in the gene or corresponding to a part of the gene not adjoining to the part of the gene corresponding to the first oligonucleotide, wherein the second oligonucleotide contains at least one ribonucleotide. A double-stranded DNA is produced using the first oligonucleotide as template and the second oligonucleotide as primer. The ribonucleotide is cleaved and removed. The nucleotide(s) located 5' to the ribonucleotide is also removed, resulting in double-stranded

fragments having cohesive ends. Such fragments are randomly reassembled by ligation to obtain novel combinations of gene sequences.

Any riboflavin biosynthesis gene or any combination of riboflavin biosynthesis genes is used for *in vitro* recombination in the context of the present invention, for example, a riboflavin biosynthesis gene derived from a plant, such as, e.g. *Arabidopsis thaliana*, e.g. a riboflavin biosynthesis gene set forth in SEQ ID NO:1 or SEQ ID NO:13, or a riboflavin biosynthesis gene from *Bacillus or E. coli*. Whole riboflavin biosynthesis genes or portions thereof are used in the context of the present invention. The library of mutated riboflavin biosynthesis genes obtained by the methods described above are cloned into appropriate expression vectors and the resulting vectors are transformed into an appropriate host, for example an algae like *Chlamydomonas*, a yeast or a bacteria. A preferred host is preferably a host that otherwise lacks riboflavin biosynthesis gene activity. Host cells transformed with the vectors comprising the library of mutated riboflavin biosynthesis genes are cultured on medium that contains inhibitory concentrations of the inhibitor and those colonies that grow in the presence of the inhibitor are selected. Colonies that grow in the presence of normally inhibitory concentrations of inhibitor are picked and purified by repeated restreaking. Their plasmids are purified and the DNA sequences of cDNA inserts from plasmids that pass this test are then determined.

An assay for identifying a modified riboflavin biosynthesis gene that is tolerant to an inhibitor may be performed in the same manner as the assay to identify inhibitors of the riboflavin biosynthesis enzyme (Inhibitor Assay, above) with the following modifications: First, a mutant riboflavin biosynthesis enzyme is substituted in one of the reaction mixtures for the wild-type riboflavin biosynthesis enzyme of the inhibitor assay. Second, an inhibitor of wild-type enzyme is present in both reaction mixtures. Third, mutated activity (activity in the presence of inhibitor and mutated enzyme) and unmutated activity (activity in the presence of inhibitor and wild-type enzyme) are compared to determine whether a significant increase in enzymatic activity is observed in the mutated activity when compared to the unmutated activity. Mutated activity is any measure of activity of the mutated enzyme while in the presence of a suitable substrate and the inhibitor. Unmutated activity is any measure of activity of the wild-type enzyme while in the presence of a suitable substrate and the inhibitor. A significant increase is defined as an increase in enzymatic activity that is larger than the margin of error inherent in the measurement technique, preferably an increase by about 2-fold or greater of the activity of the wild-type enzyme in the presence of

the inhibitor, more preferably an increase by about 5-fold or greater, most preferably an increase by about 10-fold or greater.

In addition to being used to create herbicide-tolerant plants, genes encoding herbicide tolerant riboflavin biosynthesis enzymes can also be used as selectable markers in plant cell transformation methods. For example, plants, plant tissue, plant seeds, or plant cells transformed with a transgene can also be transformed with a gene encoding an altered riboflavin biosynthesis enzyme capable of being expressed by the plant. The transformed cells are transferred to medium containing an inhibitor of the enzyme in an amount sufficient to inhibit the survivability of plant cells not expressing the modified gene, wherein only the transformed cells will survive. The method is applicable to any plant cell capable of being transformed with a modified riboflavin biosynthesis enzyme-encoding gene, and can be used with any transgene of interest. Expression of the transgene and the modified gene can be driven by the same promoter functional in plant cells, or by separate promoters.

V. Plant Transformation Technology

A wild-type or herbicide-tolerant form of the riboflavin biosynthesis gene can be incorporated in plant or bacterial cells using conventional recombinant DNA technology. Generally, this involves inserting a DNA molecule encoding the riboflavin biosynthesis enzyme into an expression system to which the DNA molecule is heterologous (i.e., not normally present) using standard cloning procedures known in the art. The vector contains the necessary elements for the transcription and translation of the inserted protein-coding sequences in a host cell containing the vector. A large number of vector systems known in the art can be used, such as plasmids, bacteriophage viruses and other modified viruses. The components of the expression system may also be modified to increase expression. For example, truncated sequences, nucleotide substitutions or other modifications may be employed. Expression systems known in the art can be used to transform virtually any crop plant cell under suitable conditions. Transformed cells can be regenerated into whole plants such that the chosen form of the riboflavin biosynthesis gene confers herbicide tolerance in the transgenic plants.

A. Requirements for Construction of Plant Expression Cassettes

Gene sequences intended for expression in transgenic plants are first assembled in expression cassettes behind a suitable promoter expressible in plants. The expression

cassettes may also comprise any futher sequences required or selected for the expression of the transgene. Such sequences include, but are not restricted to, transcription terminators, extraneous sequences to enhance expression such as introns, vital sequences, and sequences intended for the targeting of the gene product to specific organelles and cell compartments. These expression cassettes can then be easily transferred to the plant transformation vectors described *infra*. The following is a description of various components of typical expression cassettes.

1. Promoters

The selection of the promoter used in expression cassettes will determine the spatial and temporal expression pattern of the transgene in the transgenic plant. Selected promoters will express transgenes in specific cell types (such as leaf epidermal cells, mesophyll cells, root cortex cells) or in specific tissues or organs (roots, leaves or flowers, for example) and the selection will reflect the desired location of accumulation of the gene product. Alternatively, the selected promoter may drive expression of the gene under various inducing conditions. Promoters vary in their strength, i.e., ability to promote transcription. Depending upon the host cell system utilized, any one of a number of suitable promoters known in the art can be used. For example, for constitutive expression, the CaMV 35S promoter, the rice actin promoter, or the ubiquitin promoter may be used. For regulatable expression, the chemically inducible PR-1 promoter from tobacco or *Arabidopsis* may be used (*see, e.g.*, U.S. Patent No. 5,689,044).

2. Transcriptional Terminators

A variety of transcriptional terminators are available for use in expression cassettes. These are responsible for the termination of transcription beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators are those that are known to function in plants and include the CaMV 35S terminator, the *tml* terminator, the nopaline synthase terminator and the pea *rbcS* E9 terminator. These can be used in both monocotyledons and dicotyledons.

3. Sequences for the Enhancement or Regulation of Expression

Numerous sequences have been found to enhance gene expression from within the transcriptional unit and these sequences can be used in conjunction with the genes of this

invention to increase their expression in transgenic plants. For example, various intron sequences such as introns of the maize *Adh1* gene have been shown to enhance expression, particularly in monocotyledonous cells. In addition, a number of non-translated leader sequences derived from viruses are also known to enhance expression, and these are particularly effective in dicotyledonous cells.

4. Coding Sequence Optimization

The coding sequence of the selected gene may be genetically engineered by altering the coding sequence for optimal expression in the crop species of interest. Methods for modifying coding sequences to achieve optimal expression in a particular crop species are well known (see, *e.g.* Perlak *et al.*, *Proc. Natl. Acad. Sci. USA 88:* 3324 (1991); and Koziel *et al.*, *Biotechnol. 11:* 194 (1993)).

5. Targeting of the Gene Product Within the Cell

Various mechanisms for targeting gene products are known to exist in plants and the sequences controlling the functioning of these mechanisms have been characterized in some detail. For example, the targeting of gene products to the chloroplast is controlled by a signal sequence found at the amino terminal end of various proteins which is cleaved during chloroplast import to yield the mature protein (*e.g.* Comai *et al.* J. Biol. Chem. 263: 15104-15109 (1988)). Other gene products are localized to other organelles such as the mitochondrion and the peroxisome (*e.g.* Unger *et al.* Plant Molec. Biol. 13: 411-418 (1989)). The cDNAs encoding these products can also be manipulated to effect the targeting of heterologous gene products to these organelles. In addition, sequences have been characterized which cause the targeting of gene products to other cell compartments. Amino terminal sequences are responsible for targeting to the ER, the apoplast, and extracellular secretion from aleurone cells (Koehler & Ho, Plant Cell 2: 769-783 (1990)). Additionally, amino terminal sequences in conjunction with carboxy terminal sequences are responsible for vacuolar targeting of gene products (Shinshi *et al.* Plant Molec. Biol. 14: 357-368 (1990)). By the fusion of the appropriate targeting sequences described above to transgene sequences of interest it is possible to direct the transgene product to any organelle or cell compartment.

B. Construction of Plant Transformation Vectors

Numerous transformation vectors available for plant transformation are known to those of ordinary skill in the plant transformation arts, and the genes pertinent to this invention can be used in conjunction with any such vectors. The selection of vector will depend upon the preferred transformation technique and the target species for transformation. For certain target species, different antibiotic or herbicide selection markers may be preferred. Selection markers used routinely in transformation include the *nptII* gene, which confers resistance to kanamycin and related antibiotics (Messing & Vierra. Gene 19: 259-268 (1982); Bevan et al., Nature 304:184-187 (1983)), the *bar* gene, which confers resistance to the herbicide phosphinothricin (White et al., Nucl. Acids Res 18: 1062 (1990), Spencer et al. Theor. Appl. Genet 79: 625-631 (1990)), the *hph* gene, which confers resistance to the antibiotic hygromycin (Blochinger & Diggelmann, Mol Cell Biol 4: 2929-2931), and the *dhfr* gene, which confers resistance to methatrexate (Bourouis et al., EMBO J. 2(7): 1099-1104 (1983)), and the EPSPS gene, which confers resistance to glyphosate (U.S. Patent Nos. 4,940,935 and 5,188,642).

1. Vectors Suitable for *Agrobacterium* Transformation

Many vectors are available for transformation using *Agrobacterium tumefaciens*. These typically carry at least one T-DNA border sequence and include vectors such as pBIN19 (Bevan, Nucl. Acids Res. (1984)) and pXYZ. Typical vectors suitable for *Agrobacterium* transformation include the binary vectors pCIB200 and pCIB2001, as well as the binary vector pCIB10 and hygromycin selection derivatives thereof. (*See*, for example, U.S. Patent No. 5,639,949).

2. Vectors Suitable for non-*Agrobacterium* Transformation

Transformation without the use of *Agrobacterium tumefaciens* circumvents the requirement for T-DNA sequences in the chosen transformation vector and consequently vectors lacking these sequences can be utilized in addition to vectors such as the ones described above which contain T-DNA sequences. Transformation techniques that do not rely on *Agrobacterium* include transformation via particle bombardment, protoplast uptake (*e.g.* PEG and electroporation) and microinjection. The choice of vector depends largely on the preferred selection for the species being transformed. Typical vectors suitable for non-*Agrobacterium* transformation include pCIB3064, pSOG19, and pSOG35. (*See*, for example, U.S. Patent No. 5,639,949).

C. Transformation Techniques

Once the coding sequence of interest has been cloned into an expression system, it is transformed into a plant cell. Methods for transformation and regeneration of plants are well known in the art. For example, Ti plasmid vectors have been utilized for the delivery of foreign DNA, as well as direct DNA uptake, liposomes, electroporation, micro-injection, and microprojectiles. In addition, bacteria from the genus *Agrobacterium* can be utilized to transform plant cells.

Transformation techniques for dicotyledons are well known in the art and include *Agrobacterium*-based techniques and techniques that do not require *Agrobacterium*. Non-*Agrobacterium* techniques involve the uptake of exogenous genetic material directly by protoplasts or cells. This can be accomplished by PEG or electroporation mediated uptake, particle bombardment-mediated delivery, or microinjection. In each case the transformed cells are regenerated to whole plants using standard techniques known in the art.

Transformation of most monocotyledon species has now also become routine. Preferred techniques include direct gene transfer into protoplasts using PEG or electroporation techniques, particle bombardment into callus tissue, as well as *Agrobacterium*-mediated transformation.

VI. Breeding

The wild-type or altered form of a riboflavin biosynthesis gene of the present invention can be utilized to confer herbicide tolerance to a wide variety of plant cells, including those of gymnosperms; monocots, and dicots. Although the gene can be inserted into any plant cell falling within these broad classes, it is particularly useful in crop plant cells, such as rice, wheat, barley, rye, corn, potato, carrot, sweet potato, sugar beet, bean, pea, chicory, lettuce, cabbage, cauliflower, broccoli, turnip, radish, spinach, asparagus, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, quince, melon, plum, cherry, peach, nectarine, apricot, strawberry, grape, raspberry, blackberry, pineapple, avocado, papaya, mango, banana, soybean, tobacco, tomato, sorghum and sugarcane.

The high-level expression of a wild-type riboflavin biosynthesis gene and/or the expression of herbicide-tolerant forms of a riboflavin biosynthesis gene conferring herbicide tolerance in plants, in combination with other characteristics important for production and

quality, can be incorporated into plant lines through breeding approaches and techniques known in the art.

Where a herbicide tolerant riboflavin biosynthesis gene allele is obtained by direct selection in a crop plant or plant cell culture from which a crop plant can be regenerated, it is moved into commercial varieties using traditional breeding techniques to develop a herbicide tolerant crop without the need for genetically engineering the allele and transforming it into the plant.

The invention will be further described by reference to the following detailed examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless otherwise specified.

BRIEF DESCRIPTION OF THE SEQUENCES IN THE SEQUENCE LISTING

SEQ ID NO:1 is a cDNA sequence encoding the β subunit of riboflavin synthase (lumazine synthase) from *Arabidopsis thaliana*.

SEQ ID NO:2 is the predicted amino acid sequence of *Arabidopsis thaliana* lumazine synthase encoded by SEQ ID NO:1.

SEQ ID NO:3 is oligonucleotide DG-63.

SEQ ID NO:4 is oligonucleotide DG-65.

SEQ ID NO:5 is oligonucleotide JG-L.

SEQ ID NO:6 is oligonucleotide RS-1.

SEQ ID NO:7 is oligonucleotide RS-2.

SEQ ID NO:8 is a synthetic peptide used in Example 7.

SEQ ID NO:9 is a another synthetic peptide used in Example 7.

SEQ ID NO:10 is oligonucleotide DG-252.

SEQ ID NO:11 is oligonucleotide DG-253.

SEQ ID NO:12 is oligonucleotide DG-254.

SEQ ID NO:13 is a partial cDNA sequence encoding the bifunctional GTP cyclohydrolase II / DHBP synthase enzyme from *Arabidopsis thaliana*.

SEQ ID NO:14 is the predicted amino acid sequence of the mature *Arabidopsis thaliana* GTP cyclohydrolase II / DHBP synthase enzyme encoded by SEQ ID NO:13.

SEQ ID NO:15 is oligonucleotide DG-67.

SEQ ID NO:16 is oligonucleotide DG-69.

SEQ ID NO:17 is oligonucleotide DG-392a.

SEQ ID NO:18 is oligonucleotide DG-393a.

SEQ ID NO:19 is oligonucleotide DG-390a.

SEQ ID NO:20 is oligonucleotide DG-391a.

EXAMPLES

Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by Sambrook, *et al.*, Molecular Cloning, eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989) and by T.J. Silhavy, M.L. Berman, and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and by Ausubel, F.M. *et al.*, Current Protocols in Molecular Biology, pub. by Greene Publishing Assoc. and Wiley-Interscience (1987).

Example 1: Isolation of a cDNA Encoding Lumazine Synthase from *Arabidopsis*

A search of the *Arabidopsis thaliana* Expressed Sequence Tag (EST) database (Arabidopsis Biological Resource Center at Ohio State, Ohio State University, Columbus, OH) revealed an EST (EST # P25540, gb acc. # Z34233) with homology to the β Subunit of Riboflavin Synthase from *E.coli*. Using plasmid DNA of an *Arabidopsis* cDNA library (Minet *et al.*, (1992) *Plant J.* 2: 417-422) as a template, and synthetic oligonucleotides DG-63 (SEQ ID NO:3) and DG-65 (SEQ ID NO:4) designed to the EST sequence, a 204-bp DNA fragment was generated using the polymerase chain reaction (PCR). The 204-bp fragment was ligated into the TA cloning vector pCR II (Invitrogen Corp., San Diego, CA). Sequence determination by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA) confirmed that the sequence of the 204-bp fragment was identical to the sequence of EST #P25540.

Approximately150,000 pfu of a lambda ZAP *Arabidopsis* cDNA library was plated at a density of 8,000 plaques per 10 cm Petri dish, and filter lifts of the plaques were made after 7 hours growth at 37°C. The plaque lifts were probed with the 204-bp fragment labeled with 32P-dCTP by the random priming method by means of a PrimeTime kit (International Biotechnologies, Inc., New Haven, CT). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA, 1% bovine albumin at 65°C. After hybridization overnight, the filters were washed with 1% SDS, 50mM $NaPO_4$, 1mM EDTA at 65°C. Six positively hybridizing plaques were detected by autoradiography. After purification to single plaques, cDNA inserts were isolated, and their sequences were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA). A database search of the

longest clone, designated RSβ-1, using the GAP program (Deveraux, *et al.*, (1984) *Nucleic Acids Res.* 12:387-95) revealed sequence similarity to the riboflavin synthase β subunit from *E. coli*. The proteins are 68% similar and 44% identical. In addition, a comparison of the *Arabidopsis* mature protein to the *E.coli* riboflavin synthase β subunit suggests a chloroplast transit peptide is present.

RSβ-1, in the pBluescript SK vector, was deposited as pDG-4a.t. with the Agricultural Research Culture Collection (NRRL), 1815 N. University St., Peoria, IL 61604, USA under the terms of the Budapest Treaty on February 7, 1995, and assigned NRRL accession number B-21400.

The *Arabidopsis* cDNA sequence encoding RSβ-1 is set forth in SEQ ID NO:1 and the encoded amino acid sequence is set forth in SEQ ID NO:2.

Example 2: Isolation of Additional Lumazine Synthase Genes based on Sequence Similarity to the *Arabidopsis* Lumazine Synthase Coding Sequence

A phage or plasmid library is plated at a density of approximately 8,000 pfu per 10 cm Petri dish, and filter lifts of the plaques are made after 7 hours growth at 37°C. The plaque lifts are probed with the cDNA set forth in SEQ ID NO:1, labeled with 32P-dCTP by the random priming method by means of a PrimeTime kit (International Biotechnologies, Inc., New Haven, CT). Hybridization conditions are 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C. After hybridization overnight, the filters are washed with 2X SSC, 1% SDS at 50°C. Positively hybridizing plaques are detected by autoradiography. After purification to single plaques, cDNA inserts are isolated, and their sequences determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA). This experimental protocol can be used by one of ordinary skill in the art to obtain lumazine synthase genes substantially similar to the *Arabidopsis* coding sequence (SEQ ID NO:1) from any other plant species.

Example 3: Construction of a Vector Containing a GAL4 Binding Site/Minimal 35S CaMV Promoter Fused to Antisense Lumazine Synthase

pAT71:

10 GAL4 binding sites and the minimal 35S promoter (-59 to +1) were excised from pGALLuc2 (Goff, *et al.*, (1991) *Genes & Development* 5: 298-309) as an *EcoRI-PstI* fragment and inserted into the respective sites of pBluescript, yielding pAT52. pAT66 was constructed with a three-way ligation between the *HindIII-PstI* fragment of pAT52, a *PstI-EcoRI* fragment of pCIB1716 (contains a 35S untranslated leader, GUS gene, 35S terminator) and *HindIII-EcoRI* cut pUC18. The 35S leader of pAT66 was excised with PstI-NcoI and replaced with a PCR-generated 35S leader extending from +1 to +48 to yield pAT71.

pJG304:

Plasmid pBS SK+ (Stratagene, LaJolla, CA) was linearized with *SacI*, treated with mung bean nuclease to remove the *SacI* site, and re-ligated with T4 ligase to make pJG201. The 10XGAL4 consensus binding site/CaMV 35S minimal promoter/GUS gene/CaMV terminator cassette was removed from pAT71 with *KpnI* and cloned into the *KpnI* site of pJG201 to make pJG304.

pJG304 was partially digested with restriction endonuclease Asp718 to isolate a full-length linear fragment. This fragment was ligated with a molar excess of the 22 base oligonucleotide JG-L (SEQ ID NO:5). Restriction analysis was used to identify a clone with this linker inserted 5' to the GAL4 DNA binding site, and this plasmid was designated pJG304?XhoI.

pDG1:

A fragment of the lumazine synthase cDNA clone was PCR-amplified from the cDNA clone RSβ-1 using the oligonucleotides RS-1 (SEQ ID NO:6) and RS-2 (SEQ ID NO:7). This PCR product comprises the 5' portion of the lumazine synthase cDNA (SEQ ID NO:1), ending at base pair 792.

The vector pJG304?XhoI was digested with *SacI* and *NcoI* to excise the GUS gene coding sequence. The lumazine synthase PCR fragment was digested with *SacI* and *NcoI* and ligated into pJG304?XhoI to make pDG1.

Example 4: Plant Transformation Vectors For Lumazine Synthase Antisense Expression From The GAL4 Binding Site/CaMV Minimal 35S Promoter

pJG261:

Vector pGPTV (Becker, *et al.*, (1992) *Plant Molecular Biology* 20: 1195-1197) was digested with *EcoRI* and *HindIII* to remove the nopaline synthase promoter/GUS cassette. Concurrently, the superlinker was excised from pSE380 (Invitrogen, San Diego, CA) with *EcoRI* and *HindIII* and cloned into the *EcoR1/HindIII* linearized pGPTV, to make pJG261.

pDG2:

pDG1 was cut with *XhoI* to excise the cassette containing the GAL4 DNA binding site/35S minimal promoter/antisense lumazine synthase/CaMV terminator fusion. This cassette was ligated into *XhoI*-digested pJG261, such that transcription was divergent from that of the *bar* selectable marker, producing pDG2.

Example 5: Production Of GAL4 Binding Site/Minimal CaMV 35S Antisense Lumazine Synthase Transgenic Plants

pDG2 was electro-transformed (Bio-Rad Laboratories, Hercules, CA) into *Agrobacterium tumefaciens* strain C58C1 (pMP90), and *Arabidopsis* plants (Ecotype Columbia) were transformed by infiltration (Bechtold, *et al.*, (1993) *C. R. Acad. Sci. Paris*, 316: 1188-93). Seeds from the infiltrated plants were selected on germination medium (Murashige-Skoog salts at 4.3 g/liter, Mes at 0.5 g/liter, 1% sucrose, thiamine at 10 ug/liter, pyridoxine at 5 ug/liter, nicotinic acid at 5 ug/liter, myo-inositol at 1 mg/liter, pH 5.8) containing Basta at 15 mg/liter.

Example 6: Production of GAL4/C1 Transactivator Transgenic Plants

pSGZL1 was constructed by ligating the GAL4-C1 *EcoRI* fragment from pGALC1 (Goff, *et al.*, (1991) *Genes & Development*, 5: 298-309) into the *EcoRI* site of pIC20H. The

GAL4-C1 fragment of pSGZL1 was excised with *BamHI-BglII* and inserted into the *BamHI* site of pCIB770 (Rothstein, *et al.*, (1987) *Gene* 53: 153-161) yielding pAT53.

Arabidopsis root explants were transformed with pAT53 as described in Valvekens, *et al.*, (1985) *PNAS USA* 85: 5536-5540. Transgenic plants with single site insertion and positive for GAL4/C1 expression were taken to homozygosity.

Example 7: Antisense Inhibition of Lumazine Synthase Using a GAL4/C1 Transactivator and a GAL4 Binding Site/Minimal CaMV 35S Promoter

Fifteen transgenic plants containing the GAL4 binding site/minimal CaMV 35S promoter/antisense lumazine synthase construct were transplanted to soil and grown to maturity in the greenhouse. Flowers borne on the primary transformants were crossed to pollen from the homozygous GAL4/C1 transactivator line pAT53-103. F1 seeds were plated on germination medium and germination medium containing 15 mg/liter Basta. One line gave a 50% lethal phenotype on plates. Seedlings from the remaining F1 lines were transplanted to soil and grown to maturity in the greenhouse. Half of the seedlings from 2 F1 lines died while in soil.

Lumazine synthase antibody was generated in goat by injecting the synthetic peptides CIGAVIRGDTT (SEQ ID NO:8) and KAGNKGAETALTALEM (SEQ ID NO:9) conjugated to purified protein derivative. Western analysis of F1 plants revealed a significant decrease in lumazine synthase levels (Towbin *et al., PNAS USA* 76: 4350-4354).

Example 8: Expression and Purification of Recombinant Plant Lumazine Synthase in *E. coli*

To produce recombinant plant lumazine synthase in *E. coli*, a translational fusion of the *Arabidopsis* lumazine synthase cDNA (SEQ ID NO:1) to the 5' end of the thioredoxin gene (LaVallie *et al.*, (1992) *Biotechnology* 11:187-193) was created in pET-32a (Novagen, Inc., Madison, WI) using PCR. Synthetic oligonucleotide primers DG-252 (SEQ ID NO:10), DG-253 (SEQ ID NO:11), and DG-254 (SEQ ID NO:12) were used in a polymerase chain reaction to amplify DNA fragments of 693-bp and 483-bp in length. The PCR products were digested with *NcoI* and *EcoRI*. The digestion products were separated on a low-gelling-temperature agarose gel and the fragments were excised. In parallel, plasmid pET32a was digested with *NcoI* and *EcoRI*. The digestion products were separated on a gel, and the

pET32a vector was excised from the gel. The vector fragment was ligated to the two PCR-generated fragments, and the ligation products were transformed into competent *E. coli* XL1 Blue cells (Stratagene, La Jolla, CA).

Ampicillin-resistant colonies were selected, cultured, and their plasmid DNAs extracted. The structures of the plasmids were confirmed by sequencing with the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA). The recombinant plasmids with expected structure were designated pET32aRSβ FL-1 and pET32aRSβ No CTP-1.

Plasmids pET32aRSβ FL-1 and pET32aRSβ No CTP-1 were transformed into competent *E.coli* BL21(DE3) cells, and recombinant protein was expressed and purified according to the manufacturer's instructions (pET System Manual, Novagen, Inc., Madison, WI). The resulting fusion proteins produced by this strain contained approximately 132 amino acids of *E. coli* thioredoxin protein, His-Tag, and thrombin cleavage site, followed by the presumptive mature coding sequence for *Arabidopsis* lumazine synthase, which begins at codon 1 of the predicted protein coding sequence for plasmid pET32aRSβ FL-1, and codon 71 of the predicted protein coding sequence for plasmid pET32aRSβ No CTP-1.

Example 9: Lumazine Synthase Activity Assay

Lumazine synthase activity is detected using an HPLC and fluorimeter combination. Both lumazine and 2,4-dioxy-5-amino-6-ribitylamino-pyrimidine (DARP) are fluorescent under the following conditions: excitation wavelength 407 nm; emission wavelength 487 nm. However, lumazine is about 6-fold more fluorescent than an equimolar concentration of DARP. There is also a 6-fold difference in absorbance between lumazine and DARP at 405 nm. 3,4-dihydroxy-2-butanone phosphate does not fluoresce. Lumazine and DARP can be separated on a C18 column using 33% 90 mM formic acid, 60% water, and 7% methanol. Lumazine elutes first at four minutes, followed two minutes later by DARP.

The peak area can be directly related to the molar quantity of lumazine produced. Optimization studies have shown the buffer for the reaction to be preferably 100 mM KPO_4, pH 7, 5 mM β-mercaptoethanol, 2 mM DTT. The enzyme is active at a pH range of 6.5 - 7.5, but pH 7 is most preferable. Kinetic studies show that the K_m for the butanone phosphate is 190 µM and the K_m for DARP is 5.5 µM. Kis *et al.*, *Biochem.* 34: 2883-2892 (1995) reported K_m values of 130 and 5, respectively for the bacterial enzyme. The reaction

is incubated at 37°C for ten minutes and then stopped by the addition of 5% TCA. The precipitated proteins are removed by centrifugation and 10 µl of the supernatant is injected onto the HPLC. Because the reaction can proceed non-enzymatically, controls should be run with all samples to subtract this background activity.

Example 10: High Throughput Screen

A high throughput screen for novel inhibitors of lumazine synthase preferably exploits the fact that lumazine and DARP fluoresce at different intensities under optimal conditions for lumazine or the fact that there is a 6-fold difference in absorbance between these two compounds. An example of a protocol for a high throughput screen using fluorescence detection is as follows: lumazine synthase, buffer, test substance, and DARP are mixed together in the wells of a 96-well microtiter plate to a volume of 190 µl, and the initial fluorescence value is determined (with, for example, a Waters fluorimetric microtiter plate reader). Reactions commence with the addition of a 10 µl aliquot of 3,4-dihydroxy-2-butanone phosphate. After an appropriate incubation time, fluorescence is determined again. The differences between initial and final readings are then scaled as a percent of control reactions. Initial concentrations of substrates in the complete reaction mixture are preferably 50 µM for DARP and 0.5 mM for the butanone phosphate. Lumazine synthase amount and incubation time are adjusted to allow for the production of lumazine to a concentration of approximately 25 µM. This will produce a fluorescence signal that is approximately 3 to 4-fold greater than background.

Example 11: Isolation of a cDNA Encoding the Bifunctional GTP Cyclohydrolase II / 3,4-Dihydroxy-2-Butanone-4-Phosphate Synthase from *Arabidopsis*

A search of the Arabidopsis thaliana Expressed Sequence Tag (EST) database (Arabidopsis Biological Resource Center at Ohio State, Ohio State University, Columbus, OH) revealed an EST (EST # SCH1T7P; gb acc. # T12970) with homology to GTP cyclohydrolase from *Bacillus subtilis*. Using plasmid DNA of an *Arabidopsis* cDNA library (Minet et al, (1992) *Plant J.* 2: 417-422) as a template, and synthetic oligonucleotides DG-67 (SEQ ID NO:15) and DG-69 (SEQ ID NO:16) designed to the EST sequence, a 322-bp DNA fragment was generated using the polymerase chain reaction (PCR). The 322-bp

fragment was ligated into the TA cloning vector pCR II (Invitrogen Corp., San Diego, CA). Sequence determination by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA), confirmed that the sequence of the 322-bp fragment was identical to the sequence of EST # SCH1T7P.

Approximately 150,000 pfu of a lambda ZAP *Arabidopsis* cDNA library was plated at a density of 8,000 plaques per 10 cm Petri dish, and filter lifts of the plaques were made after 7 hours growth at 37°C. The plaque lifts were probed with the 322-bp fragment labeled with 32P-dCTP by the random priming method by means of a PrimeTime kit (International Biotechnologies, Inc., New Haven, CT). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA, 1% bovine albumin at 65°C. After hybridization overnight, the filters were washed with 1% SDS, 50mM NaPO4, 1mM EDTA at 65°C. Ten positively hybridizing plaques were detected by autoradiography. After purification to single plaques, cDNA inserts were isolated, and their sequences were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA). A database search of the longest clone, designated GTP-1, using the GAP program (Deveraux et al., Nucleic Acids Res. 12:387-95 (1984), revealed sequence similarity to the bifunctional GTP cyclohydrolase II/3,4-dihydroxy-2-butanone-4-phosphate synthase of *Bacillus subtilis*. The proteins are 70% similar and 54% identical. In addition, a comparison of the *Arabidopsis* mature protein to the *Bacillus subtilis* GTP cyclohydrolase II/3,4-dihydroxy-2-butanone-4-phosphate synthase suggests a chloroplast transit peptide is present.

GTP-1, in the pBluescript SK vector, was deposited as pDG-3a.t. with the Agricultural Research Culture Collection (NRRL), 1815 N. University St., Peoria, IL 61604, USA under the terms of the Budapest Treaty on February 7, 1995, and assigned NRRL accession number B-21399.

The *Arabidopsis* cDNA sequence encoding GTP-1 is set forth in SEQ ID NO:13 and the amino acid sequence of the encoded mature protein, without the putative transit peptide, is set forth in SEQ ID NO:14.

Example 12: Isolation of Additional GTP Cyclohydrolase II / 3,4-Dihydroxy-2-Butanone-4-Phosphate Synthase Genes Based On Sequence Homology the *Arabidopsis* GTP Cyclohydrolase II / 3,4-Dihydroxy-2-Butanone-4-Phosphate Synthase Coding Sequence

A phage or plasmid library is plated at a density of approximately 8,000 pfu per 10 cm Petri dish, and filter lifts of the plaques are made after 7 hours growth at 37°C. The plaque lifts are probed with the cDNA set forth in SEQ ID NO:13, labeled with 32P-dCTP by the random priming method by means of a PrimeTime kit (International Biotechnologies, Inc., New Haven, CT). Hybridization conditions are 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50°C. After hybridization overnight, the filters are washed with 2X SSC, 1% SDS at 50°C. Positively hybridizing plaques are detected by autoradiography. After purification to single plaques, cDNA inserts are isolated, and their sequences are determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA). This experimental protocol can be used by one of ordinary skill in the art to obtain bifunctional GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone-4-phosphate synthase genes substantially similar to the *Arabidopsis* coding sequence (SEQ ID NO:13) from any other plant species.

Example 13: Expression and Purification of Recombinant Plant GTP Cyclohydrolase II / DHBP Synthase in *E. coli*.

To produce recombinant higher plant GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone-4-phosphate synthase in *E.coli*, a translational fusion of the *Arabidopsis* GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone-4-phosphate synthase cDNA (SEQ ID NO:13) to the 5' end of the thioredoxin gene (LaVallie et al., *Biotechnology* 11:187-193 (1992) was created in pET-32a (Novagen, Inc., Madison, WI), using a two step PCR approach. Synthetic oligonucleotide primers DG-392a (SEQ ID NO:17) and DG-393a (SEQ ID NO:18) were used in a polymerase chanin reaction to amplify a DNA fragment of 939-bp in length. The PCR product was digested with *NcoI* and *EcoRI*. The digestion products were separated on a low-gelling-temperature agarose gel and the fragments were excised. In parallel, plasmid pET32a was digested with *NcoI* and *EcoRI*. The digestion products were separated on a gel, and the pET32a vector was excised from the gel. The vector fragment was ligated to the PCR generated fragment, and the ligation products were transformed into competent *E. coli* XL1 Blue cells (Stratagene, La Jolla, CA).

Ampicillin-resistant colonies were selected, cultured, and their plasmid DNAs extracted. The structures of the plasmids were confirmed by sequencing with the chain

termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA). The recombinant plasmid with expected structure was designated pET32aGTP-1.

Synthetic oligonucleotide primers DG-390a (SEQ ID NO:19) and DG-391a (SEQ ID NO:20) were then used in a polymerase chain reaction to amplify a DNA fragment of 662-bp. The PCR product was digested with *NcoI*. The digestion products were separated on a low-gelling-temperature agarose gel and the fragments were excised. In parallel, plasmid pET32aGTP-1 was digested with *NcoI*. The digestion products were separated on a gel, and the pET32aGTP-1 vector was excised from the gel. The vector fragment was ligated to the PCR generated fragment, and the ligation products were transformed into competent *E. coli* XL1 Blue cells (Stratagene, La Jolla, CA).

Ampicillin-resistant colonies were selected, cultured, and their plasmid DNAs extracted. The structure of the plasmids were confirmed by sequencing with the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA). The recombinant plasmid with expected structure was designated pET32aGTP-2.

Plasmid pET32aGTP-2 was transformed into competent *E. coli* BL21(DE3) cells, and recombinant protein was expressed and purified according to the manufacturer's instructions (pET System Manual, Novagen, Inc., Madison, WI). The resulting fusion proteins produced by this strain contained approximately 132 amino acids of *E. coli* thioredoxin protein, His-Tag, and thrombin cleavage site, followed by the presumptive mature coding sequence for *Arabidopsis* GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone-4-phosphate synthase.

Example 14: *In vitro* Recombination of Riboflavin Biosynthesis Genes by DNA Shuffling

A plant riboflavin biosynthesis gene (e.g., SEQ ID NO:1 or SEQ ID NO:13) encoding a riboflavin biosynthesis protein (e.g., SEQ ID NO:2 or SEQ ID NO:14, respectively) is amplified by PCR. The resulting DNA fragment is digested by DNaseI treatment essentially as described (Stemmer et al. (1994) PNAS 91: 10747-10751) and the PCR primers are removed from the reaction mixture. A PCR reaction is carried out without primers and is followed by a PCR reaction with the primers, both as described (Stemmer et al. (1994) PNAS 91: 10747-10751). The resulting DNA fragments are cloned into pTRC99a

(Pharmacia, Cat no: 27-5007-01) and transformed into a dioxygenase mutant host, e.g. by electroporation using the Biorad Gene Pulser and the manufacturer's conditions. The transformed host is grown on medium that contains inhibitory concentrations of an inhibitor selected according to a method described above, and those colonies that grow in the presence of the inhibitor are selected. Colonies that grow in the presence of normally inhibitory concentrations of inhibitor are picked and purified by repeated restreaking. Their plasmids are purified and the DNA sequences of cDNA inserts from plasmids that pass this test are then determined.

In a similar reaction, PCR-amplified DNA fragments comprising a plant riboflavin biosynthesis gene of the invention encoding a riboflavin biosynthesis protein and PCR-amplified DNA fragments comprising a riboflavin biosynthesis gene from a different host are recombined *in vitro* and resulting variants with improved tolerance to the inhibitor are recovered as described above.

Example 15: *In vitro* Recombination of Riboflavin Biosynthesis Genes by Staggered Extension Process

A plant riboflavin biosynthesis gene (e.g., SEQ ID NO:1 or SEQ ID NO:13) encoding a riboflavin biosynthesis protein (e.g., SEQ ID NO:2 or SEQ ID NO:14, respectively) and a corresponding riboflavin biosynthesis gene from a different host are each cloned into the polylinker of a pBluescript vector. A PCR reaction is carried out essentially as described (Zhao et al. (1998) Nature Biotechnology 16: 258-261) using the "reverse primer" and the "M13 20 primer" (Stratagene Catalog). Amplified PCR fragments are digested with appropriate restriction enzymes and cloned into pTRC99a and mutated riboflavin biosynthesis genes are screened as described in Example 14.

Various modifications of the invention described herein will become apparent to those skilled in the art. Such modifications are intended to fall within the scope of the appended claims.

What Is Claimed Is:

1. A DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme involved in riboflavin biosynthesis, wherein the enzyme has lumazine synthase activity or bifunctional GTP cyclohydrolase II / DHBP synthase activity.

2. A DNA molecule according to claim 1, wherein the enzyme has lumazine synthase activity.

3. A DNA molecule according to claim 2, wherein the enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:2.

4. A DNA molecule according to claim 2, wherein the enzyme comprises the amino acid sequence set forth in SEQ ID NO:2.

5. A DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme having lumazine synthase activity, wherein said DNA molecule hybridizes to a DNA molecule according to claim 4 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C.

6. A DNA molecule according to claim 2, wherein said DNA molecule hybridizes to the coding sequence set forth in SEQ ID NO:1 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C.

7. A DNA molecule according to claim 2, wherein said DNA molecule comprises a 20 base pair nucleotide portion identical in sequence to a consecutive 20 base pair portion of the coding sequence set forth in SEQ ID NO:1.

8. A DNA molecule according to claim 2, comprising the coding sequence set forth in SEQ ID NO:1.

9. A DNA molecule according to claim 1, wherein the enzyme has bifunctional GTP cyclohydrolase II / DHBP synthase activity.

10. A DNA molecule according to claim 9, wherein the enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:14.

11. A DNA molecule according to claim 9, wherein the enzyme comprises the amino acid sequence set forth in SEQ ID NO:14.

12. A DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme having bifunctional GTP cyclohydrolase II / DHBP synthase activity, wherein said DNA molecule hybridizes to a DNA molecule according to claim 11 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C.

13. A DNA molecule according to claim 9, wherein said DNA molecule hybridizes to the coding sequence set forth in SEQ ID NO:13 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C.

14. A DNA molecule according to claim 9, wherein said DNA molecule comprises a 20 base pair nucleotide portion identical in sequence to a consecutive 20 base pair portion of the coding sequence set forth in SEQ ID NO:13.

15. A DNA molecule according to claim 9, comprising the coding sequence set forth in SEQ ID NO:13.

16. A chimeric gene comprising a promoter operatively linked to a DNA molecule according to claim 1.

17. A recombinant vector comprising a chimeric gene according to claim 16, wherein said vector is capable of being stably transformed into a host cell.

18. A host cell comprising a vector according to claim 17, wherein said host cell is capable of expressing the DNA molecule encoding an enzyme involved in riboflavin biosynthesis.

19. A host cell according to claim 18, wherein said host cell is selected from the group consisting of a bacterial cell, a yeast cell, and a plant cell.

20. A host cell according to claim 19, which is a bacterial cell.

21. A process for producing nucleotides sequences encoding gene products having altered lumazine synthase activity comprising:

(a) shuffling a DNA molecule according to claim 2;

(b) expressing the resulting shuffled nucleotide sequences; and

(c) selecting for altered lumazine synthase activity as compared to the activity of an enzyme encoded by a DNA molecule according to claim 2.

22. The process of claim 21, wherein the nucleotide sequence is SEQ ID NO: 1.

23. A shuffled DNA molecule obtainable by the process of claim 22.

24. A shuffled DNA molecule according to claim 23, wherein said shuffled DNA molecule encodes an enzyme having enhanced tolerance to an inhibitor of lumazine synthase activity.

25. A chimeric gene comprising a promoter operatively linked to a shuffled DNA molecule according to claim 23.

26. A recombinant vector comprising a chimeric gene according to claim 25, wherein said vector is capable of being stably transformed into a host cell.

27. A host cell comprising a vector according to claim 26.

28. A host cell according to claim 27, wherein said host cell is selected from the group consisting of a bacterial cell, a yeast cell, and a plant cell.

29. A host cell according to claim 28, wherein said host cell is a plant cell.

30. A plant or seed comprising a plant cell according to claim 29.

31. A plant according to claim 30, wherein said plant is tolerant to an inhibitor of lumazine synthase activity.

32. A process for producing nucleotides sequences encoding gene products having altered bifunctional GTP cyclohydrolase II / DHBP synthase activity comprising:

(a) shuffling a DNA molecule according to claim 9;

(b) expressing the resulting shuffled nucleotide sequences; and

(c) selecting for altered bifunctional GTP cyclohydrolase II / DHBP synthase activity as compared to the activity of an enzyme encoded by a DNA molecule according to claim 9.

33. The process of claim 32, wherein the nucleotide sequence is SEQ ID NO: 13.

34. A shuffled DNA molecule obtainable by the process of claim 33.

35. A shuffled DNA molecule according to claim 34, wherein said shuffled DNA molecule encodes an enzyme having enhanced tolerance to an inhibitor of bifunctional GTP cyclohydrolase II / DHBP synthase activity.

36. A chimeric gene comprising a promoter operatively linked to a shuffled DNA molecule according to claim 34.

37. A recombinant vector comprising a chimeric gene according to claim 36, wherein said vector is capable of being stably transformed into a host cell.

38. A host cell comprising a vector according to claim 37.

39. A host cell according to claim 38, wherein said host cell is selected from the group consisting of a bacterial cell, a yeast cell, and a plant cell.

40. A host cell according to claim 39, wherein said host cell is a plant cell.

41. A plant or seed comprising a plant cell according to claim 40.

42. A plant according to claim 41, wherein said plant is tolerant to an inhibitor of bifunctional GTP cyclohydrolase II / DHBP synthase activity.

43. An isolated plant enzyme involved in riboflavin biosynthesis, wherein said enzyme has lumazine synthase activity or bifunctional GTP cyclohydrolase II / DHBP synthase activity.

44. An enzyme according to claim 43, wherein said enzyme has lumazine synthase activity.

45. An enzyme according to claim 44, wherein said enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:2.

46. An enzyme according to claim 44, wherein said enzyme comprises the amino acid sequence set forth in SEQ ID NO:2.

47. An enzyme according to claim 43, wherein said enzyme has bifunctional GTP cyclohydrolase II / DHBP synthase activity.

48. An enzyme according to claim 47, wherein said enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:14.

49. An enzyme according to claim 47, wherein said enzyme comprises the amino acid sequence set forth in SEQ ID NO:14.

50. A method for screening a chemical for the ability to inhibit lumazine synthase activity, comprising the steps of:

(a) combining an enzyme according to claim 44 in a first reaction mixture with 2,4-dioxy-5-amino-6-ribitylamino-pyrimidine and 3,4-dihydroxy-2-butanone phosphate under conditions in which the enzyme is capable of catalyzing the synthesis of lumazine;

(b) combining the chemical and the enzyme in a second reaction mixture with 2,4-dioxy-5-amino-6-ribitylamino-pyrimidine and 3,4-dihydroxy-2-butanone phosphate under the same conditions as in the first reaction mixture;

(c) determining the amounts of lumazine produced in the first and second reaction mixtures; and

(d) comparing the amounts of lumazine produced in the first and second reaction mixtures;

wherein the chemical is capable of inhibiting the lumazine synthase activity of the enzyme if the amount of lumazine produced in the second reaction mixture is significantly less than the amount of lumazine produced in the first reaction mixture.

51. A method according to claim 50, wherein the first reaction mixture comprises 50µM 2,4-dioxy-5-amino-6-ribitylamino-pyrimidine, and 0.5 mM 3,4-dihydroxy-2-butanone phosphate.

52. A method according to claim 50, wherein the amounts of lumazine produced in the reaction mixtures are determined using a fluorimeter at an excitation wavelength of 407 nm.

53. A chemical identified by the screening method of claim 50.

54. A method for suppressing the growth of a plant, comprising applying to the plant the chemical of claim 53, whereby the chemical inhibits the activity of lumazine synthase in the plant.

55. A method for screening a chemical for the ability to inhibit bifunctional GTP cyclohydrolase II / DHBP synthase activity, comprising the steps of:

(a) combining an enzyme according to claim 47 in a first reaction mixture with GTP or ribulose-5-phosphate under conditions in which the enzyme is capable of catalyzing the synthesis of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate, respectively;

(b) combining the chemical and the enzyme in a second reaction mixture with GTP or ribulose-5-phosphate under the same conditions as in the first reaction mixture;

(c) determining the amounts of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate produced in the first and second reaction mixtures; and

(d) comparing the amounts of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate produced in the first and second reaction mixtures;

wherein the chemical is capable of inhibiting the bifunctional GTP cyclohydrolase II / DHBP synthase activity of the enzyme if the amount of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate produced in the second reaction mixture is significantly less than the amount of 2,5-diamino-4-oxy-6-ribosylamino-pyrimidine-5'-phosphate or 3,4-dihydroxy-2-butanone phosphate produced in the first reaction mixture.

56. A chemical identified by the screening method of claim 55.

57. A method for suppressing the growth of a plant, comprising applying to the plant the chemical of claim 56, whereby the chemical inhibits the activity of GTP cyclohydrolase II / DHBP synthase in the plant.

58. A plant, plant cell, plant seed, or plant tissue comprising a DNA molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme involved in riboflavin biosynthesis, wherein the enzyme has lumazine synthase activity or bifunctional GTP cyclohydrolase II / DHBP synthase activity, and wherein the DNA molecule confers upon said plant, plant cell, plant seed, or plant tissue tolerance to a herbicide in amounts that normally inhibit riboflavin biosynthesis.

59. A plant, plant cell, plant seed, or plant tissue according to claim 58, wherein the enzyme has lumazine synthase activity, and wherein the DNA molecule confers upon the plant, plant cell, plant seed, or plant tissue tolerance to a herbicide in amounts that inhibit naturally occurring lumazine synthase activity.

60. A plant, plant cell, plant seed, or plant tissue according to claim 59, wherein the enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:2.

61. A plant, plant cell, plant seed, or plant tissue according to claim 59, wherein the DNA molecule hybridizes to the coding sequence set forth in SEQ ID NO:1 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C.

62. A plant, plant cell, plant seed, or plant tissue according to claim 58, wherein the enzyme has bifunctional GTP cyclohydrolase II / DHBP synthase activity, and wherein the DNA molecule confers upon the plant, plant cell, plant seed, or plant tissue tolerance to a herbicide in amounts that inhibit naturally occurring bifunctional GTP cyclohydrolase II / DHBP synthase activity.

63. A plant, plant cell, plant seed, or plant tissue according to claim 62, wherein the enzyme comprises an amino acid sequence substantially similar to the amino acid sequence set forth in SEQ ID NO:14.

64. A plant, plant cell, plant seed, or plant tissue according to claim 62, wherein the DNA molecule hybridizes to the coding sequence set forth in SEQ ID NO:13 under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C.

65. A method for selectively suppressing the growth of weeds in a field containing a crop of planted crop seeds or plants, comprising the steps of:

(a) planting herbicide tolerant crops or crop seeds, which are plants or plant seeds according to claim 59; and

(b) applying to the crops or crop seeds and the weeds in the field a herbicide in amounts that inhibit naturally occurring lumazine synthase activity, wherein the herbicide suppresses the growth of the weeds without significantly suppressing the growth of the crops.

66. A method for selectively suppressing the growth of weeds in a field containing a crop of planted crop seeds or plants, comprising the steps of:

(a) planting herbicide tolerant crops or crop seeds, which are plants or plant seeds according to claim 39; and

(b) applying to the crops or crop seeds and the weeds in the field a herbicide in amounts that inhibit naturally occurring bifunctional GTP cyclohydrolase II / DHBP synthase activity, wherein the herbicide suppresses the growth of the weeds without significantly suppressing the growth of the crops.